REGULATORY STATISTICS 4TH EDITION

First Revision - June 1971

BY

VICTOR C. BEAL, JR.

BIOMETRICAL SERVICES, ARS

ANIMAL HEALTH PROGRAMS

665860

CONTENTS

REGULATORY STATISTICS

QUESTIONS PART I.

1. **Define descriptive statistics.**

2. **Define inferential statistics.**

3. **List two reasons for using statistics.**

4. **What does randomization accomplish?**

5. **List three types of random sampling.**

QUESTIONS PART II.

1. What are two reasons for sampling in regulatory veterinary medicine?

   a.

   b.

2. What are four types of contamination rates?

   a.

   b.

   c.

   d.

3. Can you name other types of contamination or infection rates?

4. You wish to sample for incidence of salmonella in chickens in the United States. You take the sample as follows:

   a. Group the states according to whether they are broiler producers or egg producers. Rank them according to amount of production. Take 20% of the states in each group. Take at random 1 state from each group of 5 according to production.

   b. Take 25% of the counties in each selected state.

   c. 10% of the land area in each county. Will there be any problems? How would you do this?

   d. Every fourth chicken flock in the selected area.

   e. 10% of the chickens in the selected flock.

   Discuss the type of sampling design which you have.

QUESTIONS PART III

1. Name two types of sampling theory.

   a.

   b.

2. Calculate the probability of drawing two cards from a deck of 52 with both being hearts.

3. Name the distribution that is used in each of the two types of sampling.

   a.

   b.

4. You have a herd of 10 animals with 2 infected animals in the herd. You test 3 animals. What is the probability of all 3 tested being negative? See section 3 of Part III.

QUESTIONS PART III (CONTINUED)

5. You have 2 different herds. Each has an infection rate of 20% and you sample 40% of the animals in each herd.

Herd I has	10 animals
	2 infected animals
	4 animals sampled
Herd II has	5 animals
	1 infected animal
	2 animals sampled

   a. What is the probability of all negative animals in each sample?

      Herd I _____

      Herd II _____

   b. What happens to the probability as the herd size decreases with constant infection and sample rates?

6. You are sampling in a feed mill for Salmonella. There is a contamination rate of 10% in that feed mill. You take 5 samples. What is the probability of all 5 samples being negative? _____

7. You are sampling in a rendering plant. There is a contamination rate of 20%. You take 4 samples. What is the probability of all 4 samples being negative? _____

8. You have a herd of 800 cows which you test for Brucellosis. You wish to detect 2% infection 95% of the time. How many animals do you test?

   _____

9. If we have 100 herds in a county and wish to find out if at least 5% of them have infection, how many herds do we test with 95% probability?

   _____

10. We have a shipment of animal protein which we wish to test for Salmonella. We set a tolerance level of 20% and wish to have 95% protection. How many samples do we take? _____

QUESTIONS PART III (CONTINUED)

11. You have 1,000 shipments of dried milk, each with 5% contamination. You wish to detect 99% of the shipments or accept only 1% of them. How many samples must you take?_____

12. You have a shipment of dried milk which you wish to test for Salmonella. You set a tolerance level of 5% and a protection level of 95%. How many samples do you take? _____

13. Name the two main reasons for sampling in the rendering plant program.

14. What did the National Academy of Science say about Salmonella Free?

15. Name two types of variation.

QUESTIONS PART IV.

1. List 3 factors which affect the chance of locating infected herds with the use of Market cattle traceback.

2. We have a herd of 50 cows. We send 25 or 50% of them to slaughter over a period of 3 years. 4% of the cows are infected. 60% of the cows culled are identified through slaughter. What is the probability of detection? _____

3. We have a 100 cow herd and 40% identification through slaughter with 4% infection through slaughter with 4% infection and 50% turnover. What is the probability of detection? _____

4. We have a 100 cow herd and 40% identification rate with 4% infection and 100% turnover. What is the probability of detection? _____

5. If we have 100 herds, each with 50 cows and 10% infection and wish to detect 75 of the 100 herds when 50% of the cows are culled during a three year period, what % of the cull cows must be identified through slaughter? _____

6. Same group of herds as in 5 but we wish to detect 90 of the 100 herds. What % of the cull cows must be identified through slaughter? _____

QUESTIONS PART V.

1. What do you understand interaction to be?

2. You have the results shown from a vaccine study. Calculate the main group effects and interaction effect. Fill in missing averages.

Figure 1. Infection rates.

	45/20 present	45/20 absent	Average
Strain 19 present	0.300	0.700	
Strain 19 absent	0.600	0.900	0.750
Average	0.450		

Figure 2. Main group and interaction effects.

	45/20 present	45/20 absent	Average
Strain 19 present			
Strain 19 absent			
Average			

3. What do you understand the consequences of unequal numbers to be?

PART V CONTINUED

4. Why is it important to have proper recording of data?

5. What are some lessons to be learned from the rendering plant study?

6. What are some lessons to be learned from the diabetes paper?

REGULATORY STATISTICS PART I

INTRODUCTION

STATISTICS - THEIR USE IN ANIMAL DISEASE ERADICATION AND CONTROL

There are two fields of statistics. They are descriptive statistics and inferential statistics. Descriptive statistics is concerned with summarizing and describing data. An example of descriptive statistics is the publication entitled "Agricultural Statistics".

Inferential statistics is concerned with the analysis of data and the drawing of conclusions from this data. Inferential statistics has proven to be of value in various types of situations in the animal disease control and eradication field.

Some of the purposes for which statistics has been used include the following:

1. The design and analysis of surveys to determine the incidence of a disease.
2. The design and analysis of field trials to evaluate vaccines.
3. The design and analysis of laboratory and field studies to evaluate tuberculin.
4. The analysis of existing data in order to predict the time of eradication of various diseases.
5. To determine the probability of detecting diseased herds and flocks for different conditions of size and disease level.

A prime requirement in the interpretation of data is randomization. This is the case whether it be to determine if one vaccine is better than another or to state the incidence of a disease based on a sample of the population. Randomization in the case of a vaccine trial insures that each animal involved has an equal chance of receiving each vaccine. When only two vaccines are being compared, we can randomize by flipping a coin and assigning an animal to one vaccine if heads, and to the other vaccine if tails. Randomization in the case of a survey to determine the level of a disease insures that each animal or other sampling unit has a known chance of appearing in a survey. Random does not mean haphazard. Without proper randomization, we cannot make statements about the accuracy of a study or survey.

We will discuss some factors involved in the design of a survey. In designing a survey there are three items that are important to consider. The first is that the survey be designed so that the accuracy of the sample estimate may be determined from the sample itself providing an unbiased sample estimate. The second is that as much information as possible be obtained within the practical limits of the survey. The third is that the sample drawn be representative of the population being sampled. To obtain an unbiased sample estimate with the accuracy being determined from the sample itself, it is necessary to use some form of random sampling.

There are three types of random sampling with variations. Suppose that we have 10,000 swine that are in five states and 200 herds. Suppose that we wish to sample 200 of the swine for hog cholera. If we number the swine from 1 to 10,000 and then select 200 swine at random without regard to the state or herd origin, we have what is called simple random sampling. Suppose that we decide to select 20 of the herds at random and select the 200 swine at random from the 20 herds. We then have what is called cluster random sampling. Suppose that 50 of the herd have purebred pigs, another 50 have grade pigs, and the other 100 have crossbred pigs. Suppose that we wish to take at random 50 pigs from the purebred herds, 50 from the grade herds, and 100 from the crossbred herds. We then have what is called stratified random sampling. Suppose that we wish to select the purebred pigs from 5 herds, the grade pigs from 5 herds and the crossbred pigs from 10 herds. We then have what is called stratified cluster random sampling.

REGULATORY STATISTICS PART II

CONSIDERATIONS OF SAMPLING IN REGULATORY VETERINARY MEDICINE
(Estimation of Rates)

There are two main reasons for sampling in the field of animal disease control and eradication. The first of these is to determine estimates of rates of disease or of contamination while the second is to detect the presence of disease or of contamination. In either case, we must have an unbiased random sample.

A. Estimation of rates

We conduct sample surveys for the purpose of estimating rates. We have already discussed some of the aspects of the design of surveys in the Introduction to this booklet on Regulatory Statistics. The theory of sampling is utilized in the design and analysis of surveys. There are several good references on the theory of sampling. One of the best is Sampling Techniques by W. G. Cochran.

There are various problems in the design and conduct of surveys. One problem is that of having an accurate idea of size and makeup of the population of interest.

Several years ago when the Animal Health Division conducted a survey of Trichinae in garbage fed swine the various states had lists of the number and size of herds which were under inspection. This gave us a rather accurate idea of the population makeup and we did not encounter serious problems in the design or analysis of the survey.

More recently when the division conducted a survey on the presence of salmonella in feed, we had some serious problems in the design. While we had an estimate

of the total amount of feed produced in Basic feed mills in each state, we did not know the relative amount in each state that was cattle feed, poultry feed, swine feed, grain, plant protein, animal protein, and marine protein so that while the distribution of each of these types of material was not the same in each state, we had no choice but to have the same distribution of samples of each type in each state. This meant that while states in some areas produced mostly marine protein while states in other areas produced mostly animal protein the assignment of samples did not reflect this difference. Another problem that resulted from a lack of knowledge of the population being sampled was that we could not assign samples to feed mills on the basis of their production. This meant that mills producing large amounts of feed would not be likely to have any more samples taken then would small mills. This meant that we were not able to adjust the contamination rate according to production. Despite these problems, we were able to assure the drawing of an unbiased random sample of mills in the states that participated in the survey.

Another problem may be that of having an unbiased random sample of a large population. Since only part of the states participated in the feed survey and since these states were not chosen at random, we could not say that we had a sample that was representative of the United States. This meant that we had to say the contamination rates obtained applied only to the participating states. Another instance where we had a problem of obtaining an unbiased random sample occurred in 1968 when we wished to sample passengers coming into J. F. Kennedy Airport from overseas for the presence of agricultural material. We had to design a system of selection that was not unwieldy and yet would not yield a biased sample.

B. Types of contamination rates

There are various types of contamination rates some of which we will discuss. They include sample contamination rates, plant day contamination rate, plant contamination rate, and organism rate or organism per gram rate.

The sample contamination rate would be the proportion of samples that are positive. Of some 800 samples of animal protein in the feed survey about 30% were positive.

Another type of rate is a plant day contamination rate. At one stage of the rendering plant program about 50% of the plants would be positive on one visit. This type of rate depends upon the number of samples that are collected. If 10 samples are collected per visit more plants will have at least one positive sample than if only 5 samples are collected.

Another type of rate is a plant contamination rate. Over 80% of the plants in the rendering plant program have been positive on one or more visits. This is a plant contamination rate. This rate is affected by the number of visits and number of samples obtained.

Another rate is an organism per gram or per sample rate. This is a difficult rate to determine. It is sufficient to say that this rate was considered when the National Academy of Science Report on Salmonella was written.

REGULATORY STATISTICS PART III

CONSIDERATIONS OF SAMPLING IN REGULATORY VETERINARY MEDICINE

A. DETECTION OF DISEASE AND CONTAMINATION

1. Introduction. As was stated in Part II, a second reason for sampling in the field of animal diseases is for the detection of the presence of disease or of contamination. This subject as it relates to Market Animal Screening is covered in Part IV.

When we calculate the probability of detection of disease or of contamination, we must make use of sampling theory. We are concerned with two main types of sampling theory. The first type considered is that of sampling without replacement. We have this type when we are sampling herds of cattle. The second type is sampling with replacement. We have this type when we are sampling feed.

We make use of what is called the hypergeometric distribution when calculating probabilities in sampling without replacement. In this type of sampling the probability of a positive changes each time we test an animal. The result of one sample is not independent on the result of another sample. An illustration of this type occurs when we are dealing cards.

2. Sampling with cards. Suppose we have a deck of cards. The probability of an ace on the first draw is 1/13. The probability of an ace on the second draw is different. If we had drawn an ace on the first draw, the probability of an ace on the second draw would be 3/51, while if we did not draw an ace on the first draw the probability of an ace on the

second draw would be 4/51.

The probability of 2 aces in 2 draws = 4/52 x 3/51 = 12/2652 = .0045

The probability of no aces in 2 draws = 48/52 x 47/51 = 2256/2652 = .8507

The reason for this is that there were 52 cards before the first draw of which 48 were not aces, hence 48/52 and there were 51 cards before the second draw of which 47 were not aces if the first card was not an ace.

If we replace the card drawn after each draw, we would have sampling with replacement and would have different probabilities.

The probability of 2 aces in 2 draws = 4/52 x 4/52 = 16/2704 = .0059. Thus it can be seen that the probabilities are different when sampling with replacement.

3. Sampling herds. We will now illustrate sampling without replacement in the case of herds of cattle that are infected.

3.1 Sampling herds with 10% infection. We sample from a herd of 10 animals with 10% infection; there is 1 infected animal in the herd.

3.1.a Sampling 2 animals. If we take 2 samples the probability of all negatives is:

$$(9/10)(8/9) = 72/90 = .80$$

The probability of one positive and one negative is:

$$(1/10)(9/9) = 9/90 = .10$$

$$\text{plus } (9/10)(1/9) = 9/90 = .10$$

equals 20% probability of 1 positive in the 2 samples. We cannot have 2 positives in this case.

3.1.b Sampling 4 animals. If we take 4 samples the probability of all negatives is:

$$(9/10)(8/9)(7/8)(6/7) = 6/10 = .60$$

The 9 in 9/10 refers to the 9 negative animals and the 10 refers to the 10 total animals before taking the first sample. Each is reduced by one each time a negative animal is taken out and sampled as in the case with cards.

3.2 Sampling herds with 50% infection. In the case where we sample from a herd with 50% infection--

There are 10 animals in herd.

There are 5 infected animals.

3.2.a Sampling 2 animals. If we take 2 samples the probability of all negatives is:

$$(5/10)(4/9) = 20/90 = .2222$$

The probability of one positive and one negative is:

$$(5/10)(5/9) = 25/90 \text{ (P-N)}$$

$$\text{and } (5/10)(5/9) = 25/90 \text{ (N-P)}$$

The total (50/90) = .5556

The probability of two positives is:

$$(5/10)(4/9)=(20/90) = .2222$$

3.3 Sampling feed. We make use of what is called the binomial distribution when calculating probabilities in sampling with replacement. This is the type of situation that we assume when testing feed or rendered material for Salmonella. We will now illustrate this sampling from an

infinite population. We have this type of situation also when flipping coins or tossing dice. In sampling from an infinite population we show the various combinations of positive and negative samples with 2, 4, and 5 samples taken and incidence of positives of 10% and 50%.

3.3.a Calculation of Probability. In the case with 2 samples and 10% contamination the probability of two negatives is calculated by raising .9 to the 2nd power. This is equal to .81 as shown in Figure A. The probability of a negative and a positive is .9 times .1 or .09. The probability of a positive and then a negative is also .09. The sum of the two is .18 or the probability of getting a positive and a negative in two samples. We have shown all the possible combinations of positive and negative samples for 2 and 4 samples and part of the combinations for 5 samples. The probability of 4 negatives and one positive for 10% contamination is .9 raised to the 4th power times .1 and is equal to .06561. Since there are five ways of getting this result the total probability is 0.32805 as shown in Figure C.

In sampling from infinite population we show:

3.3.b The case of 2 samples

Figure A

	p = .10	p = .50
nn	.81	.25
nP Pn	.18	.50
PP	.01	.25

3.3.c The case of 4 samples

Figure B

	p = .10	p= .5
nnnn	.6561	.0625
nnnP nnPn nPnn Pnnn	.2916	.2500
nnPP nPnP PnnP nPPn PnPn PPnn	.0486	.3750
nPPP PnPP PPnP PPPn	.0036	.2500
PPPP	.0001	.0625

3.3.d The case of 5 samples

Figure C

	P = .10	P=.50
nnnnn	.59049	.03125
nnnnP nnnPn nnPnn nPnnn Pnnnn	.32805	.15625
nnnPP nnPnP nPnnP PnnnP nnPPn nPnPn PnnPn nPPnn PnPnn PPnnn	.07290	.31250
nnPPP	.00810	.31250
nPPPP	.00045	.15625
PPPPP	.00001	.03125

3.3.e Combinations. We can see from looking at Figure C that there are 5 ways of getting one positive sample and 10 ways of getting two positive samples. In the case of an incidence of 10%, the probability of getting one positive sample is 32.805%.

4. Sampling for detection of disease or contamination. Now that we have illustrated the principle of calculating probabilities, we will discuss sampling for the detection of disease or contamination. Sampling may or may not be appropriate. This depends upon whether or not 100% detection is desired or whether or not the attribute is not apt to exist below a certain level.

4.1 Sampling for disease. When sampling from a herd or flock for disease, we are assuming that there is a level below which the disease will not exist. We must consider the population to be a group of cattle that herd together in a bunch or to be chickens in a chicken house running freely and unseparated by pens or other types of partitions.

Tables 1, 2, and 3 show the sample sizes required for several rates of infection and flock or herd sizes ranging up to 100,000 and then to infinity.

For example, the sample size to be 95% certain of detecting 5.0% infection in a flock of chickens of size 1,000 is shown to be 57. What is the sample size for 1% infection and flock size of 5,000 for 95% Probability ?_____%.

Figures 1 and 2 summarize Table 2. It must be remembered that after we detect infection we must test all animals in order to remove the infection. In these cases we are sampling from finite populations and use the hypergeometric distribution.

4.2 Sampling for contamination. When we are sampling for Salmonella in feed, we are also sampling from a finite population but can assume an infinite population and use the binomial distribution. The reason for this is shown in Table 4. We consider a ton of feed to be divided into 1 pound, 100 gram down to 9 gram portions. For two percent contamination we have 40 contaminated units per ton in 1 pound units and must take 143 samples for detection. For 9 gram units, we have 2,015.96 contaminated units and take 148 samples for detection. It can be seen from this that we can assume infinite sampling.

The other difference from herd or flock sampling is that we can have any contamination rate down to 1/N in the population and are setting tolerance levels. This can be seen from Table 6. We have sample sizes from 5 to 458 and incidences from 60% to 1/10 percent. For a sample size of 5 we fail to detect a 60% contamination rate 1% of the time. This says that 1% of the time we will have all negative samples.

We fail to detect a 10% contamination rate 59% of the time. We see that with 90 samples that we fail to detect a 1% contamination rate 40% of the time. Thus is illustrated how tolerance levels are automatically determined from the sampling frequency.

4.2.a Reasons for sampling. There has been some controversy as to where we should sample in the rendering plant program - in-line sampling or the finished product. If one stops to think, it can be seen that we have to sample both places. The process of sampling that is being done is called quality control or acceptance sampling.

There are two reasons for this type of sampling. One is to find out where a system is out of control so that we can rectify matters and improve the product. In-line sampling is done for this purpose. We find where the contamination is occurring so that we can prevent it or reduce it by taking corrective action.

The other reason for sampling is to establish the Average Outgoing Quality Level (AOQL). The reason for this is for purposes of establishing the level of Salmonella in the product for purposes of certification. The consumer wishes to know the level of Salmonella in his product. We wish to certify plants which produce a low level Salmonella product.

Suppose that we do find contamination during the in-line sampling. We can only relate that level of contamination in the vaguest of terms to what the AOQL is. We do know that the organisms from the source of contamination will spread through the entire product. Consequently in order to evaluate final contamination, it is necessary to do finished product sampling.

4.2.b Deficiencies in detection. Finally, we must remember two things that should be emphasized over and over. That is that inadequacies in the laboratory test, plus the very nature of sampling, prevents us from

detecting all contaminated products. These two items mean that there is a built-in tolerance for Salmonella. We must also remember that valid plant to plant comparison for finished product is dependent on random sampling of outgoing product.

5. Sampling problems in meat. We have run into various problems dealing with sampling in the Animal Health Division. Some of the most difficult problems have dealt with the sampling of imported beef and pork for the thoroughness of cook and the sampling of imported horse-meat for the presence of beef. We have a problem here that there can be a very low level of defectiveness and must set a sampling rate which we feel will discourage cheating and which will not let too much uncooked meat into the country. Sampling is strictly preventive in some cases in order to assure honesty. We do not know what the minimum level of defective meat might be. In such cases such as with one shipment of hams which came into the country we were able to assume that the defective amount was equal to the contents of one cooker and set a defective rate equal to the amount of meat held by the cooker at one time divided by the total amount of meat in the shipment. The best thing to do when dealing with hams which are all one size is to require the contents of one cooker batch to be labeled and to sample one ham from each cooker batch. This would not work in the case where many cuts of meat are cooked together with a variation in the size of cuts being cooked at one time.

6. Sampling problems in ships. Another situation where sampling is not appropriate is the sampling of a portion of ships coming into the country

for forbidden ships stores. The fact that one ship from a certain country does not have forbidden items does not prevent other ships from that country from having forbidden items. A sampling of part of the ships would be setting a tolerance level where one ship with the item is dangerous.

B. CONSIDERATIONS IN DETECTING SIGNIFICANT DIFFERENCES BETWEEN GROUPS AND SIGNIFICANT CORRELATIONS BETWEEN FACTORS.

7. Types of Variation. The detection of significant differences is dependent upon the variation within groups as compared to the variation between groups. We have two types of measurements and hence two types of variation. We have continuous variation which consists of such things as pounds of milk produced by a cow, the pounds of gain in weight by an animal, the efficiency of gain or other similar measurements. We also have discrete or discontinuous variables. The testing for Salmonella gives us two outcomes which are positive or negative. Sometimes things which are discrete can appear to be close to continuous such as the number of organisms per sample. However the presence or absence is still discrete.

7.1 Further types of variation. If the measurement is positive or negative we need many more samples to detect differences between groups or between treatments than we do when the measurement is continuous. We also need more samples when we wish to detect differences among several factors such as breed of animal, type of drug, age of animal than we do when we wish to detect differences among type of drug when the animals are all one breed and age.

In experiments involving Salmonella we could get by with a smaller number of samples if all samples had some organisms in them and if we could measure the number of organisms.

8. The Use of the Chi-Square test of Significance.

8.1 The Uncorrected chi-square. This particular test is used when we wish to test for differences in proportions such as to test the difference in the proportion of animals coming down with Brucellosis when we have one group of animals that receive the standard ARS vaccine and another group of animals who receive an experimental vaccine. We will illustrate this test for just two groups of animals. When we are dealing with two groups each with two outcomes the data can be put into a two by two table. We shall show this test both with and without the use of a special correction developed by Yates. This correction is to be used only in the case of the table having two groups and two outcomes.

Figure 8.a

	Pos.	Neg.	
group a	a	b	a+b
group b	c	d	c+d
	a+c	b+d	a+b+c+d

$$X^2 = \frac{(ad-bc)^2 \ (a+b+c+d)}{(a+c) \ (b+d) \ (a+b) \ (c+d)}$$

Figure 8.b

	Pos.	Neg.	
group a	1	3	4
group b	2	2	4
	3	5	8

$$X^2 = \frac{(1 \times 2 - 3 \times 2)^2 \ 8}{(3) \ (5) \ (4) \ (4)} = .533$$

Figure 8.c

group a	5	15	20
group b	10	10	20
	15	25	40

$$X^2 = \frac{\left[(5\times10) - (15\times10)\right]^2 40}{(15)\ (25)\ (20)\ (20)} = \frac{(150 - 50)^2\ 40}{(375)\ (400)}$$

$$= \frac{(10{,}000)\ (40)}{150{,}000} = \frac{400{,}000}{150{,}000} = 2.667$$

Figure 8.d

group a	10	30	40
group b	20	20	40
	30	50	80

$$X^2 = \frac{\left[(10\times20) - (30\times20)\right]^2 80}{(30)\ (50)\ (40)\ (40)} = 5.333$$

We have illustrated the uncorrected test for the same frequency of positives and negatives but for three different sample sizes in order to show the change in the magnitude of the test of significance as the sample size increases.

8.2 The corrected chi-square. We will show the formula for the corrected test and its use for the largest of the three examples.

Figure 8.e

	positive	negative	total
Group i	a	b	a+b
Group ii	c	d	c+d
Total	a+c	b+d	a+b+c+d

$$X^2 = \frac{\left[\left|ad - bc\right| - \tfrac{1}{2}(a+b+c+d)\right]^2 (a+b+c+d)}{(a+c)\ (a+b)\ (b+d)\ (c+d)}$$

$|ad-bc|$ stands for absolute value.

For the example at the top of the page the corrected X^2 is:

$$X^2 = \frac{\left[20 \times 30 - 10 \times 20\left| - \tfrac{1}{2} 80\right]^2 (80)\right.}{30 \times 50 \times 40 \times 40} = 4.32*$$

*The order of numbers are not the same as in the lettered formula to facilitate calculation.

9. Additional Considerations in Salmonella Sampling.

9.1 Factors in detecting Salmonella. There are two factors that determine the probability of detecting Salmonella. They are the distribution (sampling distribution) of the organisms in the material being samples and the adequacy of the laboratory test. Some theoretical statements can be made about the sampling distribution. We can assume that the presence and absence of the organisms in any given sample follows the binomial distribution while the number of organisms in a sample follows the poisson distribution ranging from zero to infinity.

9.2 Tolerance levels and the NAS Report. It must always be remembered that any type of sampling for the presence of Salmonella in feed automatically sets a tolerance level. This has been pointed out in the report published by the National Academy of Science. (NAS)

9.2.a Probabilities and sample size. The NAS report made several statements. Some of them are as follows: If 60 representative 25 gram units are tested then:

1) The probability is 95% that positive units will not exceed 5% or one organism in 500 grams.
2) The probability is 99% that the lot contains less than 7% positive units or one organism per 380 grams.

3) The probability is 91% that the lot contains less than 4% positive units or one organism per 625 grams.

9.2.b Product categories. The NAS report lists 5 product categories which are as follows.

Product Category	Number of Units Tested with no positives	Unit Incidence	Significance 95% Probability of one or less organism in _____
I	60 (1500 gm)	5%	500 gm
II	29 (725 gm)	10%	250 gm
III	13 (325 gm)	20%	125 gm
IV	13 (325 gm)	20%	125 gm
V	13 (325 gm)	20%	125 gm

9.2.c Salmonella free. The report makes the statement that the term "Salmonella Free" should not be used regarding Salmonella in foods since it is not possible, with certainty to assure complete absence.

9.2.d Organisms per gram. It may be wondered the source of the concept of one organism per 625 grams or some other number of grams. I think that the answer to this is that it is assumed that the test is good enough to register a positive if there is as few as one organism in the sample.

9.2.e Organisms per gram and the poisson. We can illustrate the number of samples that might have 0, 1, 2, and so on organisms per sample under the assumption of the poisson distribution with an incidence of 1%, 5%, 10%, and 40%.

Poisson Distribution

Figure 9.a

Number of Organisms	Incidence in Units .01	.05	.10	.40
0	.99	.95	.90	.60
1	.0099498	.0487286	.0948245	.3064954
2	.0000500	.0012497	.0049954	.0782828
3	.0000002	.0000214	.0001754	.0133296
4		.0000003	.0000046	.0017023
5			.0000001	.0001739
6				.0000148
7				.0000011
8				.0000001
Positive $\bar{X}$	1.00503	1.02587	1.05361	1.27706
Total X	.01005	.05129	.10536	.51083

In the case of an incidence of 40% positive we can see that 60% of the samples would have 0 organisms, 31% would have 1 organism, 8% of the samples would have 2 organisms, etc.

Of course this would not be the case if the distribution of organisms follows some other distribution.

10. Contamination rate and organisms per sample. Some of the problems involving the sampling for Salmonella contamination have been investigated while other problems remain to be investigated. One of the problems that

has been partially investigated is that of the adequacy of the approved test. Dr. Dennis Murphy, who is the regional Poultry Epidemiologist for the North Central Region, has demonstrated that the number of organisms per sample has a definite effect upon the rate of positives. He had a range of one to approximately 100 organisms per sample. He found that when there were approximately 20 organisms per sample that 50% of the samples would be positive.

11. Effects of compositing. We need to obtain information upon the effect of compositing upon the rate of positive samples. It may also be wondered as to whether Salmonella contamination clusters or whether it is scattered randomly. There has been some work done on testing a large sample and then breaking the sample down into 10 smaller samples and testing each of those. This work was done by the Dutch. They found that they could have one or more of the sub-samples be positive when the main sample was negative.

12. Clustering of contamination. With respect to the problem of whether the contamination clusters, a small study was done in New Jersey and Pennsylvania by F. W. Germaine, who was with the Poultry Staff at that time. The results of that study indicated that there was little or no clustering although there was at times considerable variation from load to load in an individual rendering plant. Probably the amount of clustering increases when the system goes out of control.

There were six rendering plants in the study. Fifty samples were taken from each of three shipments making a total of 150 samples per plant. The results were as follows. Number of positive samples are shown.

	Shipment Number		
Plant Number	1	2	3
1	20	17	4
2	14	26	26
3	25	24	23
4	47	32	42
5	47	47	45
6	27	38	28

Tests of significance showed that there were some difference among shipments within plants for contamination rate. However, tests for randomness indicated that there were no signs of consecutive sample to be more likely to be both negative or positive as opposed to non-consecutive samples.

Table 1 -- NUMBER NEEDED TO TEST TO BE 99% CONFIDENT THAT THE DISEASE WILL BE DETECTED IF PRESENT AT OR ABOVE FIVE LEVELS OF INCIDENCE OR CONTAMINATION

FLOCK, HERD OR POPULATION SIZE = N	INCIDENCE LEVEL OR CONTAMINATION RATE = P				
	P = .1	P = .05	P = .02	P = .01	P = .005
	SAMPLE SIZE EQUALS n				
N	n	n	n	n	n
20	18	20			
50	29	41	50		
100	35	59	90	99	
150	38	67	117	143	
200	39	72	136	180	198
250	40	75	149	210	244
300	41	77	159	235	286
350	↓	79	167	255	324
400	41	80	174	272	360
450	42	81	179	287	391
500	↓	82	183	300	420
600	↓	83	189	320	470
700	42	84	194	336	511
800	43	85	198	349	546
1,000	↓	86	204	367	601
1,200	↓	86	208	381	642
1,400	↓	87	210	391	673
1,600	↓	87	212	398	699
1,800	↓	88	214	404	719
2,000	↓	↓	215	409	736
3,000	43	88	219	425	791
4,000	44	89	222	433	821
5,000	↓	↓	223	438	839
6,000	↓	↓	224	441	852
10,000	↓	89	225	448	888
100,000	↓	90	228	457	915
00	44	90	228	458	919

Table 2 -- NUMBER NEEDED TO TEST TO BE 95% CONFIDENT THAT THE DISEASE WILL BE DETECTED IF PRESENT AT OR ABOVE FIVE LEVELS OF INCIDENCE OR CONTAMINATION

FLOCK, HERD OR POPULATION SIZE = N	INCIDENCE LEVEL OR CONTAMINATION RATE = P				
	P = .10	P = .05	P = .02	P = .01	P = .005
	SAMPLE SIZE EQUALS n				
N	n	n	n	n	n
20	15	19			
50	22	34	48		
100	25	44	77	95	
150	26	48	94	129	
200	27	51	105	155	190
250	↓	52	112	174	
300		53	117	189	
350		54	121	201	
400	27	54	124	210	310
450	28	55	126	218	
500	↓	55	128	224	
600		56	131	235	378
700		↓	134	243	
800		56	135	249	421
1,000	28	57	138	258	450
1,200	↓	↓	139	264	471
1,400			141	268	486
1,600			142	272	500
1,800		57	142	275	509
2,000		58	143	278	517
3,000		↓	145	284	542
4,000			146	287	555
5,000			↓	289	563
6,000			146	291	569
10,000			147	294	580
100,000			148	298	596
∞	28	58	148	298	598

Table 3 -- NUMBER NEEDED TO TEST TO BE 90% CONFIDENT THAT THE DISEASE WILL BE DETECTED IF PRESENT AT OR ABOVE FIVE LEVELS OF INCIDENCE OR CONTAMINATION

FLOCK, HERD OR POPULATION SIZE = N	INCIDENCE LEVEL OR CONTAMINATION RATE = P				
	P = .10	P = .05	P = .02	P = .01	P = .005
	SAMPLE SIZE EQUALS n				
N	n	n	n	n	n
20	13	18	20	20	20
50	18	30	45	50	50
100	20	36	68	90	100
150	20	39	80	118	150
200	21	40	87	136	180
250		41	92	150	210
300		42	95	160	235
350			97	168	256
400		42	99	174	273
450		43	101	180	288
500			102	184	300
600	21		104	190	321
700	22	43	105	195	337
800		44	106	199	349
1,000			108	205	368
1,200			109	209	382
1,400			109	211	392
1,600			110	213	399
1,800			110	215	405
2,000		44	111	216	410
3,000		45	112	221	426
4,000			112	223	434
5,000			113	224	439
6,000				225	442
10,000			113	227	449
100,000			114	229	458
00	22	45	114	229	459

Table 4 -- NUMBER OF SAMPLES REQUIRED FOR VARYING LEVELS OF CONTAMINATION AND VARYING SAMPLE SIZES PER TON OF MATERIAL

	1 Pound	100 Grams	50 Grams	18 Grams	9 Grams
Sample Units Per Ton	2,000	9,073	18,144	50,399	100,798
10% Contamination					
Number of Contaminated Units Per Ton	200	907.3	1,814.4	5,039.9	10,079.8
Number of Samples	28	28	28	28	28
Pounds Destroyed	28	6.17	3.09	1.11	0.56
2% Contamination					
Number of Contaminated Units Per Ton	40	181.46	362.88	1,007.98	2,015.96
Number of Samples	143	147	148	148	148
Pounds Destroyed	143	32.41	16.31	5.87	2.94
1/2% Contamination					
Number of Contaminated Units Per Ton	10	45.365	90.72	251.995	503.99
Number of Samples	517	578	588	594	596
Pounds Destroyed	517	127.43	64.82	23.57	11.83
Number of Samples Per Pound	1	4.5359	9.0718	25.2	50.4

Table 5 -- PROBABILITY OF FAILING TO DETECT THE CONTAMINATED LOT

LOT PERCENT CONTAMINATED	n=5	n-10	n-13	n-28	n=36	n=58	n-148	n=298	n=30
1%	95	90	88	75	70	56	23	5	
2%	90	82	77	57	48	31	5	0.2	
3%	86	74	67	43	33	17	1	0.01	
4%	82	66	59	32	23	9	0.2	0	
5%	77	60	51	24	16	5	0.05	0	
6%	73	54	45	18	11	3	0.01	0	
7%	70	48	39	13	7	1	0	0	
8%	66	43	34	10	5	0.8	0	0	
9%	62	39	29	7	3	0.4	0	0	
10%	59	35	25	5	2	0.2	0	0	
12%	53	28	19	3	1	0.06	0	0	
14%	47	22	14	1	0.4	0.02	0	0	
16%	42	17	10	0.8	0.2	0	0	0	
18%	37	14	8	0.4	0.08	0	0	0	
20%	33	11	5	0.2	0.03	0	0	0	
26%		5							
37%		1							
45%	5								
60%	1								

Table 6 -- PROBABILITY OF FAILING TO DETECT THE CONTAMINATED LOT FOR GIVEN INCIDENCES AND SAMPLE SIZE

Sample Size	5	10	13	21	28	44	58	90	298	458
Estimated Incidence			.20	.20	.10	.10	.05	.05	.01	.01
Probability			.05	.01	.05	.01	.05	.01	.05	.01
PROBABILITY OF FAILING TO DETECT THE CONTAMINATED LOT FOR ACTUAL INCIDENCE AND ABOVE SAMPLE SIZE										
Actual Incidence										
.60	.01	$1.0X10^{-4}$	$6.7X10^{-6}$	$4.4X10^{-9}$						
.45	.05	$2.5X10^{-3}$	$4.2X10^{-4}$	$3.5X10^{-6}$	$5.4X10^{-8}$					
.37	.10	.01	$2.5X10^{-3}$	$6.1X10^{-5}$	$2.4X10^{-6}$	$1.5X10^{-9}$				
.30	.17	.03	.01	$5.6X10^{-4}$	$4.6X10^{-5}$	$1.5X10^{-7}$	$1.0X10^{-9}$			
.26	.22	.05	.02	$1.8X10^{-3}$	2.2X10-4	$1.8X10^{-6}$	2.6X10-8			
.25	.24	.06	.02	$2.4X10^{-3}$	$3.2X10^{-3}$	$3.2X10^{-6}$	5.7X10-8			
.20	.33	.11	.05	.01	$1.9X10^{-3}$	$5.4X10^{-5}$	$2.4X10^{-6}$	$1.9X10^{-9}$		
.15	.44	.20	.12	.03	.01	$7.8X10^{-4}$	$8.1X10^{-5}$	$4.4X10^{-7}$		
.10	.59	.35	.25	.11	.05	.01	$2.2X10^{-3}$	$7.6X10^{-5}$		
.09	.62	.39	.29	.14	.07	.02	$4.2X10^{-3}$	$2.1X10^{-4}$		
.08	.66	.43	.34	.17	.10	.03	$7.9X10^{-3}$	$5.5X10^{-4}$		
.07	.70	.48	.39	.22	.13	.04	.01	$1.5X10^{-3}$	$4X10^{-10}$	
.06	.73	.54	.45	.27	.18	.07	.03	$3.8X10^{-3}$	$9.8X10^{-9}$	
.05	.77	.60	.51	.34	.24	.10	.05	.01	$2.3X10^{-7}$	$1X10^{-10}$
.04	.82	.66	.59	.42	.32	.17	.09	.03	$5.2X10^{-6}$	$7.6X10^{-9}$
.03	.86	.74	.67	.53	.43	.26	.17	.06	$1.1X10^{-4}$	$8.7X10^{-7}$
.02	.90	.82	.77	.65	.57	.41	.31	.16	$2.4X10^{-3}$	$9.6X10^{-5}$
.01	.95	.90	.88	.81	.75	.64	.56	.40	.05	.01
.008	.96	.92	.90	.84	.80	.70	.63	.49	.09	.03
.006	.97	.94	.92	.88	.84	.77	.71	.58	.17	.06
.004	.98	.96	.95	.92	.89	.84	.79	.70	.30	.16
.002	.99	.98	.97	.96	.95	.92	.89	.84	.55	.40
.001	.995	.99	.99	.98	.97	.96	.94	.91	.74	.63

Fig.-1 95 % PROBABILITY GRAPH FOR FIVE INCIDENCE RATES FOR POPULATIONS VARYING FROM 10 to 100,000

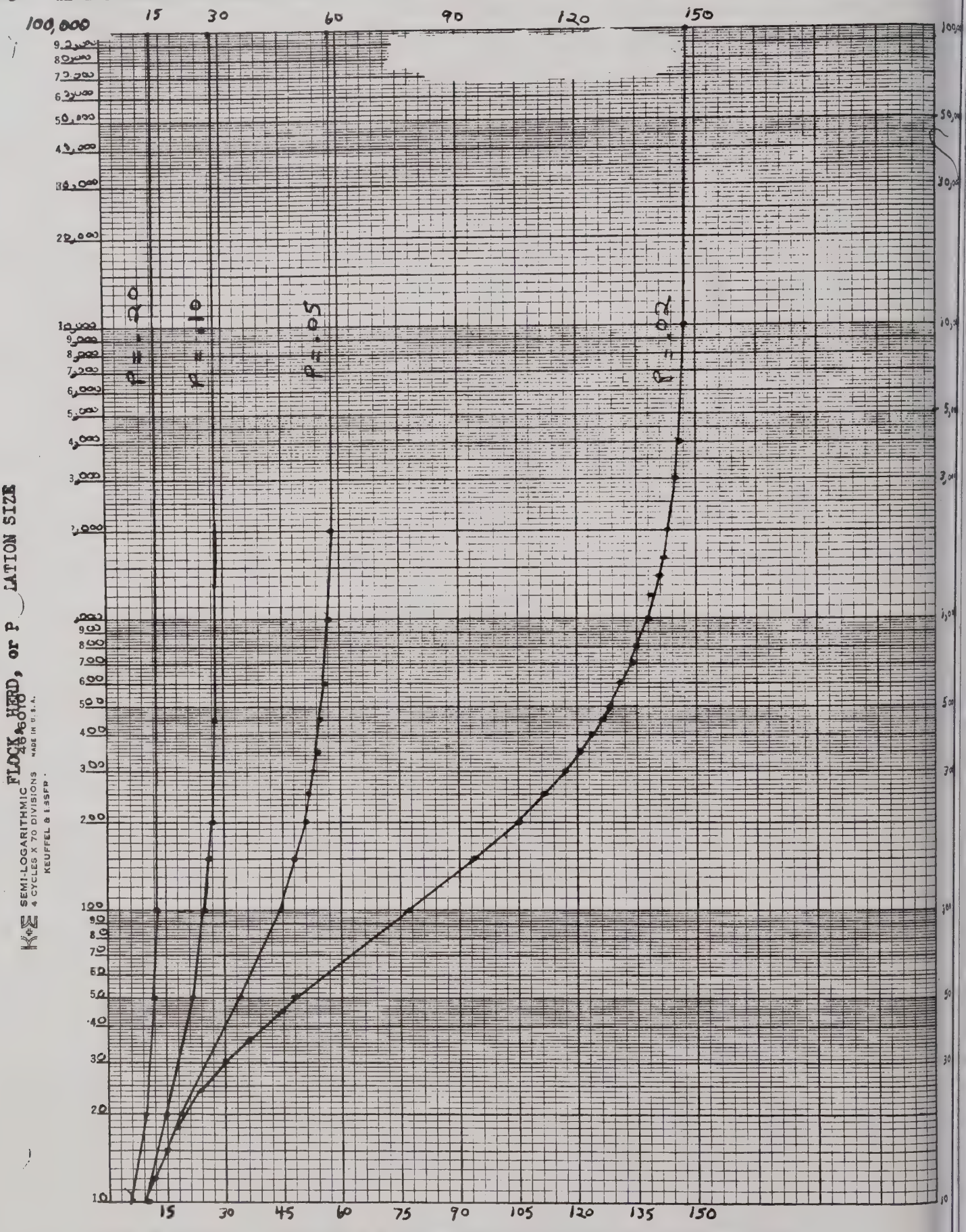
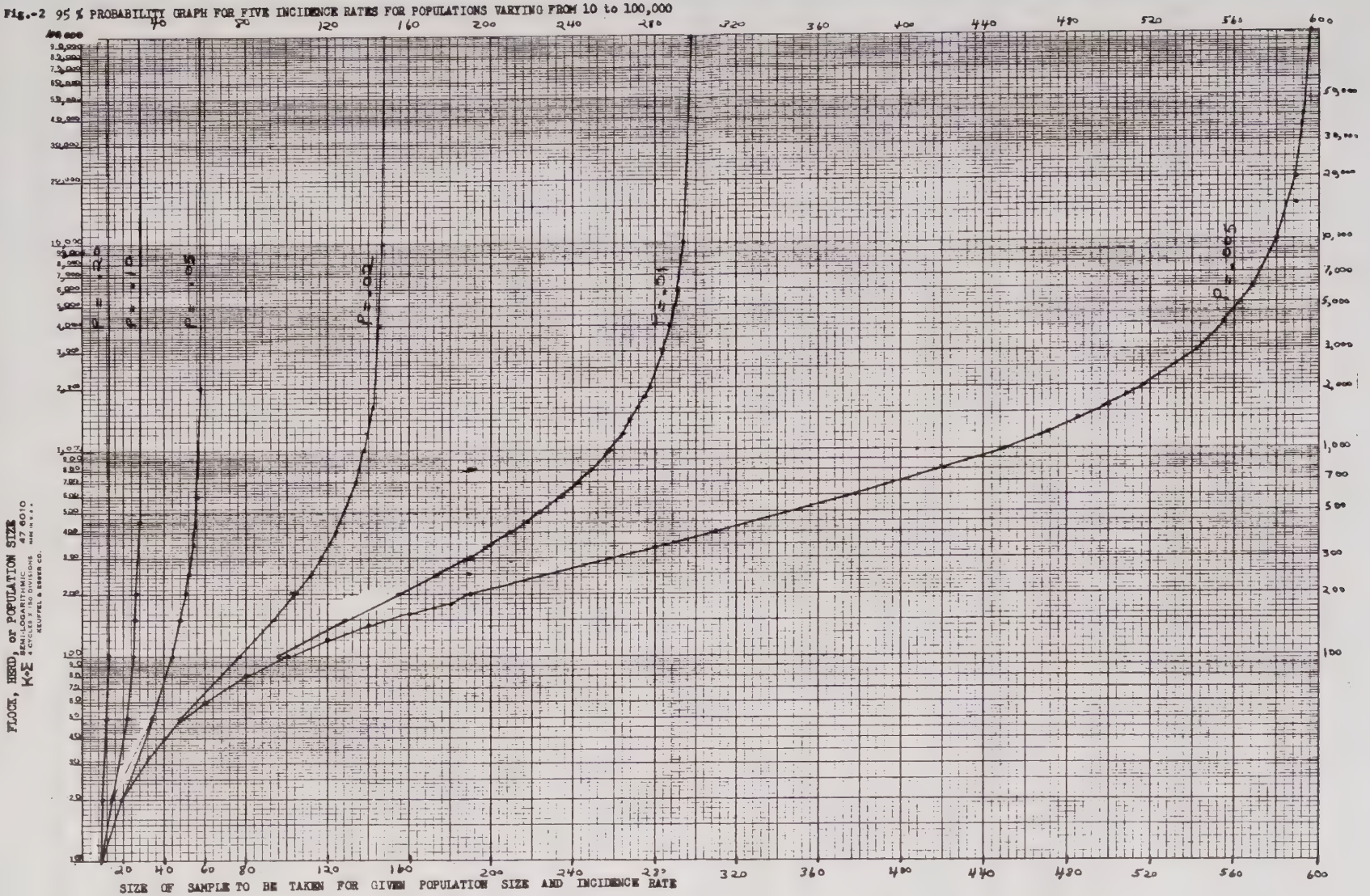


Fig.-2 95 % PROBABILITY GRAPH FOR FIVE INCIDENCE RATES FOR POPULATIONS VARYING FROM 10 to 100,000



REGULATORY STATISTICS -- APPENDIX PART III

ESTIMATION OF SAMPLE SIZE IN THE DETECTION OF DISEASED POPULATIONS

1. Introduction. The prime need in the field of Animal Disease Control and Eradication work is to detect herds or flocks that are infected. This is a simple task when the disease is manifested clinically and is under continuous surveillance. However, with many diseases there must be blood tests, serological tests, or other tests to detect the presence of infection.

In many cases it is too expensive to test all animals in a herd or flock either because of the expense of the test or because of the expense of handling the animals. In such cases it is necessary to sample a proportion of the herd or flock. In order to do this, we must decide the size of the sample which should be taken. The size of sample is determined by the population size, infection rate which is to be detected, and the percent[probability or (1 - a) level where a represents the error that we are willing to accept, such as a = 5%,] of such infected units which are to be detected. The hypergeometric distribution is used for calculation. Once a herd or flock is determined to be infected, all animals are tested.

2. Historical. Dr. Walter Harvey, formerly of Biometrical Services, first worked on this problem for the Animal Disease Eradication Division about 1957. Originally, the desired sample sizes were computed by the process of iteration with a computer. This process limited the combinations of population size, infection rate, and (1 - a) values for which sample sizes could be reasonably computed. Further, it is difficult to answer the question as to what infection rate will be detected with a certain probability for a specific rate of sampling.

Values provided by Dr. Harvey were used to construct graphs for 2% infection and (1 - a) values of 95% and 99%. The graph for 95% was inserted into the UNIFORM METHODS AND RULES FOR BRUCELLOSIS.

In 1963 the question was asked of the author as to what sample size would be required to detect infection rates in the order of one-half percent. After examination of the terms of the hypergeometric distribution, it occurred to the author that there might be an approximation that would be easier to work with.

It did develop that there is a relatively simple approximation which may be used in the computing of sample sizes for specific infection rates and values of (1 - a) as well as computing the infection rate which would be detected for specific sample sizes and (1 - a) values.

The question remained as to the accuracy of the estimates. The first comparisons were made with the estimates provided by Dr. Harvey and with the actual a values which might be calculated using the actual hypergeometric distribution and relatively small numbers. After the ANH Division acquired a Wang 380 desk calculator, it became possible to calculate the a values using Stirling's approximation to the factorial.

The hypergeometric approximation provides sample size estimates which have an accuracy of greater than 99% in the value of a for most of the infection rates, population sizes, and (1 - a) values which are of interest. This approximation has been of considerable value in the field of animal disease work.

3. Definitions.

   a. N = Total animals in the herd.

   b. d = Number of infected animals in the herd.

c. $(N - d)$ = Number of infection-free animals in the herd.

d. n = Number of animals tested or examined, i.e., the sample.

e. $(N - n)$ = Number of animals not tested, i.e., not in the sample.

f. r = Number of reactors = number of infected animals in the sample.

g. $N!$ = N factorial = $N(N - 1)(N - 2).....(3)(2)(1)$.

h. $4!$ = 4 factorial = $4 \times 3 \times 2 \times 1 = 24$.

i. $0!$ = zero factorial = 1 by definition.

j. $\binom{5}{3}$ = 5 things taken 3 at a time = $\frac{5!}{3!\,(5 - 3)!} = \frac{(5 \times 4 \times 3 \times 2 \times 1)}{(3 \times 2 \times 1)\,(2 \times 1)}$.

k. $\binom{N}{d}$ = N things taken $\underline{d}$ at a time as above. = $\frac{N!}{d!\,(N - d)!}$

l. a_o = The probability of not detecting the disease = the probability of not having an infected animal in the sample.

m. $Ln(x)$ = The natural logarithm of the number of value (x).

n. i_{th} = the i_{th} number or individual or item.

o. a_i = T he probability of having $\underline{i}$ infected individuals in the sample, i.e., $a_{(i = 1)}$ = the probability of 1 infected individual in the sample.

4. The Use of the Hypergeometric Distribution. As mentioned above and in Part III, the hypergeometric distribution is used in the calculation of probabilities for the detection of disease in finite populations. The distribution is shown below using the symbolism in section 3 above.

$$a_i = \frac{\binom{d}{r = i}\binom{N - d}{n - r}}{\binom{N}{n}}$$

$$= \frac{\frac{d!}{r!\,(d - r)!}\,\frac{(N - d)!}{(n - r)!\,(N - d - n + r)!}}{\frac{N!}{(N - n)!\,n!}} \qquad (4.1)$$

$$a_o = \frac{\binom{d}{0}\binom{N-d}{n-0}}{\binom{N}{n}}$$

$$= \frac{\dfrac{d!}{0!\,(d-0)!}\;\dfrac{(N-d)!}{(n-0)!\,(N-d-n+0)!}}{\dfrac{N!}{(N-n)!\,n!}} \qquad (4.2)$$

$$= \frac{\dfrac{d!\;(N-d)!}{d!\;n!\;(N-d-n)!}}{\dfrac{N!}{n!\;(N-n)!}}$$

$$= \frac{d!\;(N-d)!\;(N-n)!\;n!}{d!\;n!\;(N-d-n)!\;N!}$$

$$= \frac{(N-d)!\;(N-n)!}{(N-d-n)!\;N!} \qquad (4.3)$$

$$= \frac{(N-n)\;(N-n-1)\text{------}(N-n-d+2)\;(N-n-d+1)}{N(N-1)\;(N-2)\text{----------}(N-d+2)\;(N-d+1)} \left[\frac{(N-d-n)!\,(N-d)!}{(N-d)!\;(N-d-n)!}\right]$$

The portion in brackets equals one and cancels leaving the portion in parentheses.

4.1 A Numerical Example of the Hypergeometric Distribution. Let:

$N = 20;\ d = 4;\ n = 10$

$$a_o = \frac{\binom{4}{0}\binom{16}{10}}{\binom{20}{10}} = \frac{\dfrac{4!}{0!\;4!}\;\dfrac{16!}{10!\;6!}}{\dfrac{20!}{10!\;10!}} = \frac{16!\;10!\;10!}{10!\;6!\;20!}$$

$$= \frac{16!\;10!}{6!\;20!}$$

$$= \frac{16!\;10 \times 9 \times 8 \times 7 \times 6!}{6!\;20 \times 19 \times 18 \times 17 \times 16!}$$

$$= \frac{10 \times 9 \times 8 \times 7}{20 \times 19 \times 18 \times 17} = 0.04334$$

5. The Approximation to the Hypergeometric. We shall now develop an approximation to the hypergeometric which is relatively simple to work with when we desire to estimate sample size.

$$a_o = \frac{(N-d)!\ (N-n)!}{N!\ (N-n-d)!} = \frac{(N-n)\ (N-n-1)\text{------}(N-n-d+2)\ (N-n-d+1)}{N(N-1)\ (N-2)\ (N-3)\text{-----}(N-d+2)\ (N-d+1)} \qquad (4.3)$$

$$\simeq \frac{\left\{\frac{(N-n)\ +\ (N-n-d+1)}{2}\right\}^d}{\left\{\frac{N\ +\ (N-d+1)}{2}\right\}^d} \simeq \frac{(N\ -\ n\ -\ d/2\ +\ \tfrac{1}{2})^d}{(N\ -\ d/2\ +\ \tfrac{1}{2})^d} \qquad (5.1)$$

5.1 We then take the logarithm of both sides of the equation.

$$\text{Ln}\ (a_o) = d\ \ \text{LN}(N-n-d/2\ +\ \tfrac{1}{2})\ -\ d\ \ \text{LN}(N\ -\ d/2\ +\ \tfrac{1}{2}) \qquad (5.2)$$

$$d\ \text{Ln}(N\ -\ n\ -\ d/2\ +\ \tfrac{1}{2}) = d\ \text{Ln}(N\ -\ d/2\ +\ \tfrac{1}{2})\ +\ \text{Ln}\ (a_o) \qquad (5.3)$$

5.2 We wish to obtain a sample size (n) that will give us the probability (1-α) of a certain size of detecting disease. Consequently we substitute x for $(N - n - d/2 + \tfrac{1}{2})$ in the above equation so that we have:

$$\text{Ln}\ (a_o) + d\ \text{Ln}(N\ -\ d/2\ +\ \tfrac{1}{2}) = d\ \text{Ln}(x) \qquad (5.4)$$

We then solve for Ln(x),

$$\text{Ln}(x) = \text{Ln}(a_o)/d + d\ \text{Ln}(N\ -\ d/2\ +\tfrac{1}{2})/d \qquad (5.5)$$

We then take the anti-log to obtain the value of (x) and solve for n.

$$x = N\ -\ n\ -\ d/2\ +\ \tfrac{1}{2} \qquad (5.6)$$

$$n = N\ -\ d/2\ +\ \tfrac{1}{2}\ -\ x \qquad (5.7)$$

5.3 We may also make use of the approximation to solve for the number of infected animals (d) in the herd for which we might expect to get at least one infected animal in a specific sample size with a specific probability level. For this purpose the approximation takes the following form:

$$a_o \cong \frac{(N - d - n/2 + \frac{1}{2})^n}{(N - n/2 + \frac{1}{2})^n} \tag{5.8}$$

$$\text{Ln}\left(a_o\right) = n\text{Ln}(N - d - n/2 + \tfrac{1}{2}) - n\ \text{Ln}(N - n/2 + \tfrac{1}{2}) \tag{5.9}$$

We then solve for d in the same manner that we solved for n in 5.2.

5.4 Conditions for Greatest Accuracy. The greatest amount of accuracy of estimation is obtained when solving for sample size (n) if the value of n is large in relation to the value of d. When solving for the number of infected animals (d) the value of d should be large in relation to the value of n.

6. Investigation of the Accuracy of the Approximation to the Hypergeometric. It is difficult to work with factorials when they become very large. It is also difficult to work with them when any of the values are other than whole numbers. In this situation, we can use Stirling's approximation to the factorial which is extremely accurate where--

$$N! = (2\ pi)^{\frac{1}{2}}\ N^{N + \frac{1}{2}}\ e^{-N + 1/12N - 1/360\ N^3 + 1/1260\ N^5 - \ldots} \tag{6.1}$$

where pi = 3.1416 and e = 2.71828.

We do not need the last two terms in the exponent of e to obtain sufficient accuracy when using the Stirling approximation in the hypergeometric.

We take the logarithm of the factorial giving us as follows:

$$\text{Ln}(N!) = (N + \tfrac{1}{2})\ \text{Ln}(N) - N + \tfrac{1}{2}\ \text{Ln}(2\ pi) + 1/(12\ N) - 1/360N^3 \tag{6.2}$$

We substitute the above terms into the formula for the hypergeometric:

$$\text{Ln}(a) = \text{Ln}(N\text{-}d)! + \text{Ln}(N\text{-}n)! - \text{Ln}(N)! - \text{Ln}(N\text{-}n\text{-}d)! \tag{6.3}$$

The value of n which is used is that obtained in the approximation in 5.2:

$$a = \frac{(N-d)^{N-d+\frac{1}{2}}\,(N-n)^{N-n+\frac{1}{2}}\,e^{1/12(N-d)}\;e^{1/12(N-n)}\;e^{-1/360(N-d)^3}\;e^{-1/360(N-n)^3}}{(N)^{N+\frac{1}{2}}\,(N-n-d)^{N-n-d+\frac{1}{2}}\,e^{1/12(N)}\;e^{1/12(N-n-d)}\;e^{-1/360N^3}\;e^{-1/360(N-n-d)^3}} \qquad (6.4)$$

We take the logarithm of the above equations and have the following:

$$\begin{aligned} Ln(a) = {} & (N-d+\tfrac{1}{2})\ln(N-d) + (N-n+\tfrac{1}{2})\,Ln(N-n) \\ & - (N+\tfrac{1}{2})\,Ln\,N - (N-n-d+\tfrac{1}{2})\,Ln\,(N-n-d) \\ & + \frac{1}{12\,(N-d)} + \frac{1}{12\,(N-n)} - \frac{1}{12\,N} - \frac{1}{12\,(N-n-d)} \\ & - \frac{1}{360(N-d)^3} - \frac{1}{360(N-n)^3} + \frac{1}{360(N)^3} + \frac{1}{360(N-n-d)^3} \end{aligned} \qquad (6.5)$$

6.1 The Approximation Gives Sufficient Accuracy. In the actual calculation of the values of a, the last term using $1/360N^3$ was not used since there was accuracy to several places without it. With most populations sizes for infection rates of 20% or less and alpha levels of 1% to 10%, the actual alpha level is less than the approximated alpha level for given sample sizes. Conversely, the estimated sample size for a given alpha level is greater than the required sample size; however, the margin of error in alpha is generally less than 1%. When we have certain sampling rates and infection rates, the margin of error in alpha is greatest when we use as the exponent the larger number in using the hypergeometric approximation. Concersely, if we are estimating either the sample rate or the infection rate for given population sizes, and alpha levels for the other rate, the rate which is estimated with greatest accuracy is the larger of the two rates.

APPENDIX III A

Definitions and Examples of Detection Rates for Arbitrary Contamination Rates and Distribution of Contamination.

The probabilities shown in the Tables in PART III are for the case where all sampling is done with complete randomization of a certain population. This restriction applies for a single days sampling of a rendering plant or of a herd of cattle or a flock of chickens. When we are talking about the entire years production of a rendering plant where we are taking 30 samples we must take 30 different days at random for the given probabilities to apply.

In the actual situation that exists where we are taking 10 samples on each of three different days we may talk about the probability of detection for each individual day. In order to talk about the probability of detection over the period of a year we must know what the variation in contamination from day to day is for individual plants. This is due to the fact that we have restricted the sampling to 3 days or clusters. Each of the days during the year is a cluster with 3 of the clusters being selected. We know very little as to what manner the contamination rate varies from day to day in individual plants. However we can invent arbitrary examples to illustrate how much the probability of detection might vary over the period of a year.

Five examples have been invented. Two of the examples are intended to illustrate a continuous range of contamination.

The 1st of these examples has three contamination rates, each occuring on 1/3 of the days during the year. The arbitrary contamination rates are;

A = .05
B = .10
C = .15

This provides an overall contamination rate of .10.

Table 3 shows the distribution of the three contamination rates along with the probabilities of having all samples be negative when 10 samples are obtained. Table 4 shows the distribution of the possible sampling outcomes. The total probability of failure to detect equal 0.05559.

The second of the examples has five contamination rates. Table 5 shows the various contamination rates with their frequency, f(p), and the probability, q^{10}, of having all negative samples when 10 samples are obtained. Table 6 shows the distribution of all possible sampling outcomes along with the average contamination rate (P), frequency of outcome, $f(\bar{q})$, $(q = 1 - p)$, the product of the failure to detect for the various combinations ($\Pi\, q_i^{10} = q_a^{10}\, q_b^{10}\, q_c^{10}$ for the combination ABC), and this product times the frequency of the outcome. The probability of failure to detect for this example is 0.0681. The probability of obtaining one or more positive samples is 0.9319.

The other three examples are very arbitrary in that they assume that the plant has a constant contamination rate on part of the days during the year and that on the other days there is a zero contamination rate. This is an unnatural situation but the examples serve to illustrate the fact that with a given contamination rate that the probabilities of detection can vary. These examples make use of the Double Binomial Distribution which is shwon on page 5 of Part one. This is a very good distribution for describing the probability of obtaining brucellosis or tuberculosis reactors where:

p describes the infection rate within a herd;

$q = 1 - p$; r describes the probability of an infected animal having a positive reaction;

s = 1 - r;

m = the number of tests performed on an individual animal and n = the number of animals tested. The formula is $[p\ (r + s)^{m} + q]^{n}$. The probability of having all negative tests = $(q + ps^{m})^{n}$.

In the three examples shown here:

(p) = the proportion of days during the year when there is contamination in the plant;

q represents the proportion of days when there is no contamination;

r represents the contamination rate on the contaminated days;

s represents 1 - r;

m represents the number of samples obtained on a sampling day;

and n represents the number of days sampled.

The example shown in Table shows the results when the contamination rate is 10 percent and values are of p, q, r, s, m, and n are as shown below

p	1.00	0.50	0.40	0.25	0.20	0.10
q	0.00	0.50	0.60	0.75	0.80	0.90
r	0.10	0.20	0.25	0.40	0.50	1.00
s	0.90	0.80	0.75	0.60	0.50	0.00

m	1	2	3	5	6	10	15	30
n	30	15	10	6	5	3	2	1
mn	30	30	30	30	30	30	30	30

It can be seen from the body of the table that the lowest probability of failure occurs for the case of 1 sample on each of 30 days and for the case of contamination rate (r) = 10% on each day of the year (p = 1.00).

It can also be seen that the highest probabilities of failure occur in the case of 30 samples on 1 day and the case of a contamination rate of 100% on 10% of the days with the highest probability of failure occuring on the combination of the two cases. There is a probability of failure of 0.2413 = 24.13% for the case of 10 samples on 3 visits and a contamination rate of 25% on 40% of the days. This particular example may be fairly reasonable in describing what might actually happen to the contamination rate from day to day when the overall rate is 10%.

The other two examples are shown in Tables 8 and 9. These two table describe the same sampling frequencies as Table 7, however Table 8 describes an overall rate of 20% while Table 9 describes an overall rate of 25%.

TABLE 1 Probabilities of Not Detecting with 30 Samples and Varying Contamination Rates

Contamination Rate P	Q	$(Q)^{30}$	Q	$(Q)^{30}$
.01	.99	.7397	.81	.00180
.02	.98	.5455	.80	.00124
.03	.97	.4010	.79	.00085
.04	.96	.2939	.78	.00058
.05	.95	.2146	.77	.00039
	.94	.1563	.76	.00027
	.93	.1134	.75	.00018
	.92	.0820	.74	.00012
	.91	.0591	.73	.00008
	.90	.0424	.72	.00005
	.89	.0303	.71	.00003
	.88	.0216	.70	.00002
	.87	.0153	.69	.0000150
	.86	.0108	.68	.000009
	.85	.00763	.67	.000006
	.84	.00535	.66	.000004
	.83	.00374	.65	.000002
	.82	.00260		

P = Contamination Rate

Q = Rate not Contaminated

P + Q = 1

Q^{30} = Q raised to 30th Power

= Probability 30 Consecutive Negative Samples given Q

TABLE 2 Probabilities of Not Detecting with 10 Samples and Varying Contamination Rates

Contamination Rate P	Q	$(Q)^{10}$	Q	$(Q)^{10}$
.01	.99	0.904382	.75	0.056314
	.98	0.817073	.74	0.049240
	.97	0.737424	.73	0.042976
	.96	0.664833	.72	0.037439
	.95	0.598737	.71	0.032552
	.94	0.538615	.70	0.028248
	.93	0.483982	.69	0.024462
	.92	0.434388	.68	0.021139
	.91	0.389416	.67	0.018228
	.90	0.348678	.66	0.015683
	.89	0.311817	.65	0.013463
	.88	0.278501	.64	0.011529
	.87	0.248423	.63	0.009849
	.86	0.221302	.62	0.008393
	.85	0.196874	.61	0.007133
	.84	0.174901	.60	0.006047
	.83	0.155160	.59	0.005111
	.82	0.137448	.58	0.004308
	.81	0.121577	.57	0.003620
	.80	0.107374	.56	0.003033
	.79	0.094683	.55	0.002533
	.78	0.083358	.54	0.002108
	.77	0.073267	.53	0.001749
	.76	0.064289	.52	0.001446
	.75	0.056314	.51	0.001190
			.50	0.000977

P = Contaminated Rate

Q = Rate not Contaminated

P + Q = 1

Q^{10} = Q raised to 10th Power

= Probability 10 Consecutive Negative Samples given Q

TABLE 3 Example of a Population Having Three Contamination Rates

	P	Q	f(p)	pf(p)	Q^{10}
A	.05	.95	1/3	0.01667	0.598737
B	.10	.90	1/3	0.0333	0.348678
C	.15	.85	1/3	0.0500	0.196874
			1.00	0.1000	

P = Contamination Rate

Q = Rate not Contaminated

P + Q = 1

A, B, C = 3 different rates of contamination on days of the year

F (P) = proportion of days with each rate

Total of PF (P) = Overall contamination rate

Q^{10} = probability all 10 samples being negative

TABLE 4 Continuation of Example 1

The Disbribution of all Possible Sampling Results

Combination	$\bar{P}$	$\bar{Q}$	$f(\bar{q})$	$\Pi\, q_i^{10}$	$\Pi\, q_i^{10}\, f(\bar{q})$
AAA	0.05	0.95	1/27	0.215	0.00796
AAB	0.067	0.933		0.125	0.00463
ABA	"	"		"	"
BAA	"	"		"	"
ABB	0.083	0.917		0.073	0.00270
BAB	"	"		"	"
BBA	"	"		"	"
BBB	0.1	0.9		0.042	0.00156
AAC	0.083	0.917		0.071	0.00263
ACA	"	"		"	"
CAA	"	"		"	"
ACC	0.117	0.883		0.023	0.00085
CAC	"	"		"	"
CCA	"	"		"	"
CCC	0.15	0.85		0.008	0.00030
3BBC	0.117	0.883	3/27	0.024	0.00267
3BCC	0.133	0.867	3/27	0.014	0.00157
6ABC	0.1	0.9	6/27	0.041	0.00911
			1.00		0.05559

AAA etc = possible combinations of contaminations from Table 3.

$\bar{P}$ = Average contamination rate over the 3 sampling days.

$\bar{Q} = 1-\bar{P}$

$f(\bar{q})$ = Proportion of possible samplings having respative combination of contamination rates.
Example AAA = $\left(\frac{1}{3}\right)\left(\frac{1}{3}\right)\left(\frac{1}{3}\right) = 1/27$

$\Pi\, q_i^{10}$ = Product of 10 values from Table 3
Example for ABA = (.599) (.349) (.599) = .125

Overall probability of Detection = Total of $\Pi\, q_i^{10}\, f(\bar{q})$ column.

TABLE 5 Example of a Population Having Five Contamination Rates

	P	f(p)	pf(p)	q^{10}
A	.00	.07	.0000	1.0000
B	.05	.29	.0145	0.598737
C	.10	.32	.0320	0.348678
D	.15	.21	.0315	0.196874
E	.20	.11	.0220	0.107374
		1.00	0.1000	

P = Contamination rate.

Q = Rate not Contaminated.

P + Q = 1

A, B, C, D, E = 5 different rates of contamination on days of year.

f (p) = Proportion of days with each rate.

Total of pf(p) = Overall contamination rate.

Q^{10} = Probability of 10 samples being negative.

TABLE 6 Distribution of Results for 3 Plant Visits and 5 Outcomes*

	$\bar{P}$	$\bar{Q}$	$f(\bar{q})$	πq_i^{10}	$\pi q_i^{10} f(\bar{q})$
A^3	0.0000	1.0000	0.000343	1.00000	0.000343
$3A^2B$	0.0167	0.9833	0.004263	0.598737	0.002552
$3A^2C$	0.0333	0.9667	0.004704	0.348678	0.001640
$3A^2D$	0.05	0.95	0.003087	0.196874	0.000608
$3A^2E$	0.0667	0.9333	0.001617	0.107374	0.000174
$3AB^2$	0.0333	0.9667	0.017661	0.358486	0.006331
6ABC	0.05	0.95	0.038976	0.208767	0.008137
6ABD	0.0667	0.9333	0.025578	0.117876	0.003015
6ABE	0.0833	0.9167	0.013398	0.064289	0.000861
$3AC^2$	0.0667	0.9333	0.021504	0.121577	0.002614
6ACD	0.0833	0.9167	0.028224	0.068646	0.001937
6ACE	0.1	0.9	0.014784	0.037439	0.000553
$3AD^2$	0.1	0.9	0.009261	0.038760	0.000359
6ADE	0.1167	0.8833	0.009702	0.021139	0.000205
$3AE^2$	0.1333	0.8667	0.002541	0.011529	0.000029
B^3	0.05	0.95	0.024389	0.214639	0.005235
$3B^2C$	0.0667	0.9333	0.080736	0.124996	0.010092
$3B^2D$	0.0833	0.9167	0.052983	0.070577	0.003739
$3B^2E$	0.1	0.9	0.027753	0.038492	0.001068
$3BC^2$	0.0833	0.9167	0.089088	0.072792	0.006485
6BCD	0.1	0.9	0.116928	0.041101	0.004806
6BCE	0.1167	0.8833	0.061248	0.022416	0.001373
$3BD^2$	0.1167	0.8833	0.038367	0.023207	0.000890
6BDE	0.1333	0.8667	0.040194	0.012657	0.000509
$3BE^2$	0.15	0.85	0.010527	0.006903	0.000073
C^3	0.1	0.9	0.032768	0.042391	0.001389
$3C^2D$	0.1167	0.8833	0.064512	0.023935	0.001544
$3C^2E$	0.1333	0.8667	0.033792	0.013054	0.000441
$3CD^2$	0.1333	0.8667	0.042336	0.013515	0.000572
6CDE	0.15	0.85	0.044352	0.007371	0.000327
$3CE^2$	0.1667	0.8333	0.011616	0.004020	0.000047
D^3	0.15	0.85	0.009261	0.007631	0.000071

* Columns defined in Table 4.

TABLE 6 (cont)

	$\bar{p}$	$\bar{q}$	$f(\bar{q})$	$\prod q_i^{10}$	$\prod q_i^{10} f(\bar{q})$
$3D^2E$	0.1667	0.8333	0.014553	0.004162	0.000061
$3DE^2$	0.1833	0.8167	0.007623	0.002270	0.000017
E^3	0.2	0.8	0.001331	0.001238	0.000002
					0.068100

Table 7. Probabilities of failure to detect with overall contamination rate of 10% and varying values of p, q, r, s, m, and n.

			q	0.00	0.5	0.6	0.75	0.80	0.90
			s	0.90	0.8	0.75	0.60	0.50	0.00
			p	1.00	0.50	0.40	0.25	0.20	0.10
			r	0.10	0.20	0.25	0.40	0.50	1.00
m	n	mr	pr	0.10	0.10	0.10	0.10	0.10	0.10
1	30	30		0.0424	0.0424	0.0424	0.0424	0.0424	0.0424
2	15	30		0.0424	0.0510	0.0558	0.0731	0.0874	0.2059
3	10	30		0.0424	0.0610	0.0721	0.1129	0.1461	0.3487
5	6	30		0.0424	0.0856	0.1126	0.2075	0.2747	0.5314
6	5	30		0.0424	0.1001	0.1362	0.2563	0.3341	0.5905
10	3	30		0.0424	0.1697	0.2413	0.4244	0.5124	0.7290
15	2	30		0.0424	0.2679	0.3664	0.5627	0.6400	0.8100
30	1	30		0.0424	0.5006	0.6001	0.7500	0.8000	0.9000

m = number of samples taken on individual day

n = number of days upon which samples are taken

mn = total number of samples = 30

q = proportion of days plant product is negative

p = proportion of days plant product is contaminated

p + q = 1

s = rate of product not contaminated

r = contamination rate on days in which there is contamination

r + s = 1

pr = overall contamination rate = .10

Values in body of Table = Probability All samples being negative.

TABLE 8 Probabilities of failure to detect with overall contamination rate of 20% and varying values of p,q, r, s, m and n

			q	0.00	0.20	.50	..60	0.75	0.80
			s	0.80	0.75	.60	.50	0.20	0.00
			p	1.00	0.80	.50	.40	0.25	0.20
			r	0.20	0.25	.40	.50	0.80	1.00
m	n	mn\pr		0.20	0.20	0.20	0.20	0.20	0.20
1	30	30		0.0012	0.0012	0.0012	0.0012	0.0012	0.0012
2	15	30		0.0012	0.0016	0.0031	0.0047	0.016	0.035
3	10	30		0.0012	0.0020	0.0069	0.013	0.058	0.107
5	6	30		0.0012	0.0035	0.024	0.053	0.178	0.262
6	5	30		0.0012	0.0047	0.039	0.082	0.237	0.328
10	3	30		0.0012	0.015	0.127	0.216	0.422	0.512
15	2	30		0.0012	0.044	0.250	0.360	0.5625	0.640
30	1	30		0.0012	0.200	0.500	0.600	0.75	0.800

m = number of samples taken on individual day

n = number of days upon which samples are taken

mn = total number of samples = 30

q = proportion of days plant product is negative

p = proportion of days plant product is contaminated

p + q = 1

s = rate of product not contaminated

r = contamination rate on days in which there is contamination

r + s = 1

pr overall contamination rate = .20

Values in body of Table = Probability All samples being negative.

TABLE 9 Probabilities of failure to detect with overall contamination rate of 25% and varying values of p, q, r, s, m and n.

			q	0.00	0.375	0.50	0.600	.6875	0.75
			s	0.75	0.600	0.50	0.375	.200	0.00
			p	1.00	0.625	.50	.400	.3125	.25
			r	0.25	0.400	.50	.625	.800	1.00
m	n	mn \ pr		.25	.25	.25	.25	.25	.25
1	30	30		0.00018	0.00018	0.00018	0.00018	0.00018	0.00018
2	15	30		0.00018	0.00047	0.00087	0.00180	0.0047	0.013
3	10	30		0.00018	0.00119	0.00317	0.00854	0.024	0.056
5	6	30		0.00018	0.0058	0.01879	0.04806	0.106	0.178
6	5	30		0.00018	0.0108	0.03377	0.07848	0.154	0.237
10	3	30		0.00018	0.0543	0.12537	0.216	0.325	0.422
15	2	30		0.00018	0.141	0.25002	0.360	0.473	0.56
30	1	30		0.00018	0.375	0.50000	0.600	0.6875	0.75

m = number of samples taken on individual day

n = number of days upon which samples are taken

mn = total number of samples = 30

q = proportion of days plant product is negative

p = proportion of days plant product is contaminated

p + q = 1

s = rate of product not contaminated

r = contamination rate on days in which there is contamination

r + s = 1

pr overall contamination rate = .25

Values in body of Table = Probability All samples being negative.

REGULATORY STATISTICS PART IV

CONSIDERATIONS ON DETECTING DISEASED HERDS WITH MARKET CATTLE IDENTIFICATION

INTRODUCTION

Market cattle identification and traceback are important parts of the Tuberculosis and Brucellosis Eradication programs. Various factors affect the chance of locating infected herds with the use of Market Cattle Traceback (MCT). They include: (1) the number of animals culled and sent to slaughter; (2) the size of the herd; (3) the amount of infection in the herd; (4) the rate of MCT coverage of cull animals; and (5) the marketing patterns or chance of cull animals from individual herds being identified and going to slaughter plants where blood samples are collected and/or where satisfactory examinations for evidence of tuberculosis with carcass identification are conducted.

The tables and graphs presented in this section examine the effect of changes in all but the last of the above variables. The actual probabilities are lower than those shown because marketing patterns are not random. It is shown that low culling rates, small herd sizes, low infection rates, and low MCT coverage decrease the chance of finding infected herds. In beef herds having 20 cows with a 10 percent infection rate and 1/6 turnover per year, a 40 percent MCT sampling coverage will result in 40 percent of the infected herds being found over a three-year period, while a coverage of 80 percent will result in 80 percent of the herds being found. In beef herds with 50 cows and the same infection and turnover rates, 40 percent MCT coverage will result in 72 percent of the infected herds being found, while 80 percent MCT

coverage will result in 98 percent of the infected herds being found.

DISCUSSION

The tables and graphs shown here were originally prepared for the tuberculosis program and appeared in the 1965 Tuberculosis Committee report to the United States Livestock Sanitary Association. However, the percentages shown apply to other diseases where the principles of traceback are utilized. The probabilities are based upon the assumption of random marketing patterns and do not consider the effect of some farmers and ranchers being likely to sell their cull animals in markets where there is no identification or sending them to plants where there is no slaughter inspection or bleeding. The actual probabilities are lower than those shown due to the lack of random marketing.

Calculations on the probabilities of finding infected herds under various conditions are based on:

(1) Rate of herd turnover for slaughter.

The probability figures in tables one and three are based on a 50 percent turnover during a three-year period. This more or less coincides with the turnover in beef herds.

Tables two and four are based on a 100 percent herd turnover during a three-year period. This corresponds more closely to management practices in dairy herds.

(2) Size of herd.

(3) Potential herd infection rates - 10, 5, 4, or 2

percent, based on animals with lesions (blood titers).

(4) Maintenance of the same rate of infection during the entire period.

Tables one and two show the probability of finding infected herds under three rates of animal identification at slaughter -- 40, 60, and 80 percent.

Table one is based on a 50 percent herd turnover to slaughter during a three-year period. It shows that in a 10-cow herd with a 10 percent herd infection and with a 40 percent animal identification through slaughter, there is a 20 percent probability of finding disease in three years. In a 20-cow herd having a 10 percent infection and 40 percent animal identification through slaughter, there is a 40 percent probability of finding disease. It also shows that in the 20-cow herd with only a 5 percent infection, the probability of finding the disease is only 40 percent when 80 percent of the animals are identified. The average herd size in many parts of the country does not exceed 20 head of cows.

Table two is based on 100 percent herd turnover to slaughter during a three-year period. This increases the probability index. In tables one and two, it is apparent that with a stated rate of herd infection, the probability of finding the disease is greater in herds with high infection rates than in herds with low infection rates.

It is also apparent that infected animals are more readily revealed in herds having the higher rates of animal movements to slaughter.

Tables three and four show the animal identification rate that is necessary to find a specific percentage of the infected herds. The probabilities

specified in table three are based on a 50 percent herd turnover during a three-year period. In table four, the probabilities specified are based on a 100 percent herd turnover during a three-year period.

Herd infection rate and herd size are again basic factors in determining the probability of finding infected herds. For instance, table three shows that in a 10-cow herd with a 10 percent infection during a three-year period, there is a 50 percent chance of finding infection on slaughter if 100 percent of the cattle are identified. In a 50-cow herd with a 10 percent infection, there is a 75 percent chance of finding the infection if 44 percent of the cattle are identified on slaughter.

The probabilities shown in tables three and four indicate that a higher rate of identification coverage is needed than now exists to have an effective eradication program based on post-mortem examination or blood samples with traceback to herds of origin.

Figures one, two, and three summarize some of the information from tables one, two, three, and four. They are based on 10 percent infection. Figure one summarizes tables one and two. The left-hand set of bars shows that for 40 percent MCT coverage and 50 percent turnover, the percent of infected herds identified decreases, as the size of herds decreases from 100 to 10 cows. Figure two summarizes tables three and four. The bars on the left illustrate that it is necessary to increase the level of sampling or examination of identified slaughter cattle as the herd size decreases to locate the same percentage of infected herds. For example, at a 10 percent animal infection rate within each herd, and 50 percent turnover of animals within each herd over a three-year period, it would be necessary to sample 12 cows

from each 100-cow herd to locate 75 percent of the infected herds; 11 cows from each 50-cow herd, and 8 cows from each 20-cow herd to maintain this same rate of locating infected herds. Comparison with the set of bars for location of 95 percent of the infected herds shows that the required rate of MCT coverage must increase in order to locate a greater percentage of the infected herds.

The significance of these relationships can be best illustrated by the needs of each of our programs. For example, the Uniform Methods and Rules for Tuberculosis requires that five percent of the adult animals per year be identified for reaccreditation and 10 percent per year for free status. This amounts to 15 percent and 30 percent, respectively, over a three-year period. When there is a turnover of about 16 percent per year, or 48 percent over a three-year period, this amounts to an identification rate of about 31 percent and 62 percent, respectively. The tables show the consequences when the identification and sampling rate is only 40 percent. Therefore, we can anticipate even lower levels of effectiveness when the rate approaches only 31 percent.

Even an identification rate through slaughter of 60 percent does not insure that more than 30 percent of the infected herds in the case of five percent infection and 20 cow herds will be located. The actual situation with the chance of farmers selling all of their cull animals into channels where identification does not occur is even worse. It can be seen that we need an identification rate close to 100 percent for small herds, or herds with low infection rates. The same logic applies to other diseases.

TABLE 1 -- PROBABILITY OF FINDING INFECTED HERDS DURING A 3-YEAR PERIOD UNDER THE MARKET CATTLE IDENTIFICATION PROGRAM (BASED ON A 50% HERD TURNOVER DURING THE 3-YEAR PERIOD)

CATTLE IN HERD	HERD INFECTION RATE	ANIMALS MOVED TO SLAUGHTER DURING A 3-YEAR PERIOD	LESION OR TITERED ANIMALS MOVED TO SLAUGHTER DURING A 3-YEAR PERIOD	PROBABILITY OF FINDING DISEASE WHEN ANIMAL IDENTIFICATION RATE THROUGH SLAUGHTER IS --		
				40 PERCENT	60 PERCENT	80 PERCENT
NUMBER	PERCENT	NUMBER	NUMBER	PERCENT	PERCENT	PERCENT
10 ---	10	5	1/2	20	30	40
20 ---	10	10	1	40	60	80
	5	10	1/2	20	30	40
50 ---	10	25	2 1/2	72	89	98
	4	25	1	40	60	80
	2	25	1/2	20	30	40
100 ---	10	50	5	92	98	99
	4	50	2	64	84	96
	2	50	1	40	60	80

TABLE 2 -- PROBABILITY OF FINDING INFECTED HERDS DURING A 3-YEAR PERIOD UNDER THE MARKET CATTLE IDENTIFICATION PROGRAM (BASED ON 100% HERD TURNOVER DURING THE 3-YEAR PERIOD)

CATTLE IN HERD	HERD INFECTION RATE	ANIMALS MOVED TO SLAUGHTER DURING A 3-YEAR PERIOD	LESION OR TITERED ANIMALS MOVED TO SLAUGHTER DURING A 3-YEAR PERIOD	PROBABILITY OF FINDING DISEASE WHEN ANIMAL IDENTIFICATION RATE THROUGH SLAUGHTER IS ------		
				40 PERCENT	60 PERCENT	80 PERCENT
NUMBER	PERCENT	NUMBER	NUMBER	PERCENT	PERCENT	PERCENT
10 ---	10	10	1	40	60	80
20 ---	10	20	2	64	84	96
	5	20	1	39	60	80
50 ---	10	50	5	92	98	99
	4	50	2	64	84	96
	2	50	1	40	60	80
100 ---	10	100	10	99.6	100	100
	4	100	4	88	98	99.9
	2	100	2	64	84	96

TABLE 3 -- ANIMAL RATE THROUGH SLAUGHTER REQUIRED TO FIND INFECTED HERDS DURING A 3-YEAR PERIOD UNDER THE MARKET CATTLE IDENTIFICATION PROGRAM (BASED ON A 50% HERD TURNOVER DURING THE 3-YEAR PERIOD)

CATTLE IN HERD	HERD INFECTION RATE	ANIMALS MOVED TO SLAUGHTER DURING A 3-YEAR PERIOD	LESION OR TITERED ANIMALS MOVED TO SLAUGHTER DURING A 3-YEAR PERIOD	ANIMAL IDENTIFICATION RATE REQUIRED ON SLAUGHTER WHEN THE DESIRED PROBABILITY OF FINDING INFECTION IS --			
				50 PERCENT	75 PERCENT	90 PERCENT	95 PERCENT
NUMBER	PERCENT	NUMBER	NUMBER	PERCENT	PERCENT	PERCENT	PERCENT
10 ---	10	5	1/2	100	*	*	*
20 ---	10	10	1	**	80	90	100
	5	10	1/2	100	*	*	*
50 ---	10	25	2 1/2	**	44	60	68
	4	25	1	**	76	92	96
	2	25	1/2	100	*	*	*
100 ---	10	50	5	**	24	36	44
	4	50	2	**	50	68	78
	2	50	1	**	76	90	96

* 50% PROBABILITY OF DETECTING DISEASE IS BEST THAT CAN BE ACHIEVED WITH THESE INFECTION AND TURNOVER RATES.

TABLE 4 -- ANIMAL IDENTIFICATION RATE REQUIRED TO FIND INFECTED HERDS DURING A 3-YEAR PERIOD UNDER THE MARKET CATTLE IDENTIFICATION PROGRAM (BASED ON 100% HERD TURNOVER DURING THE 3-YEAR PERIOD)

CATTLE IN HERD	HERD INFECTION RATE	ANIMALS MOVED TO SLAUGHTER DURING A 3-YEAR PERIOD	LESION OR TITERED ANIMALS MOVED TO SLAUGHTER DURING A 3-YEAR PERIOD	ANIMAL IDENTIFICATION RATE REQUIRED ON SLAUGHTER WHEN THE DESIRED PROBABILITY OF FINDING INFECTION IS-		
				75 PERCENT	90 PERCENT	95 PERCENT
NUMBER	PERCENT	NUMBER	NUMBER	PERCENT	PERCENT	PERCENT
10 ---	10	10	1	80	90	100
20 ---	10	20	2	50	70	80
	5	20	1	75	90	95
50	10	50	5	24	36	44
	4	50	2	50	68	78
	2	50	1	76	90	96
100	10	100	10	12	20	25
	4	100	4	29	43	52
	2	100	2	50	68	77

FIGURE 1 ---- RELATIONSHIP OF CULL RATES, HERD SIZES, ANIMAL INFECTION RATE AND PERCENT OF ANIMALS IDENTIFIED, EXAMINED, AND SAMPLED DURING SLAUGHTER WITH THE PERCENT OF INFECTED HERDS FOUND

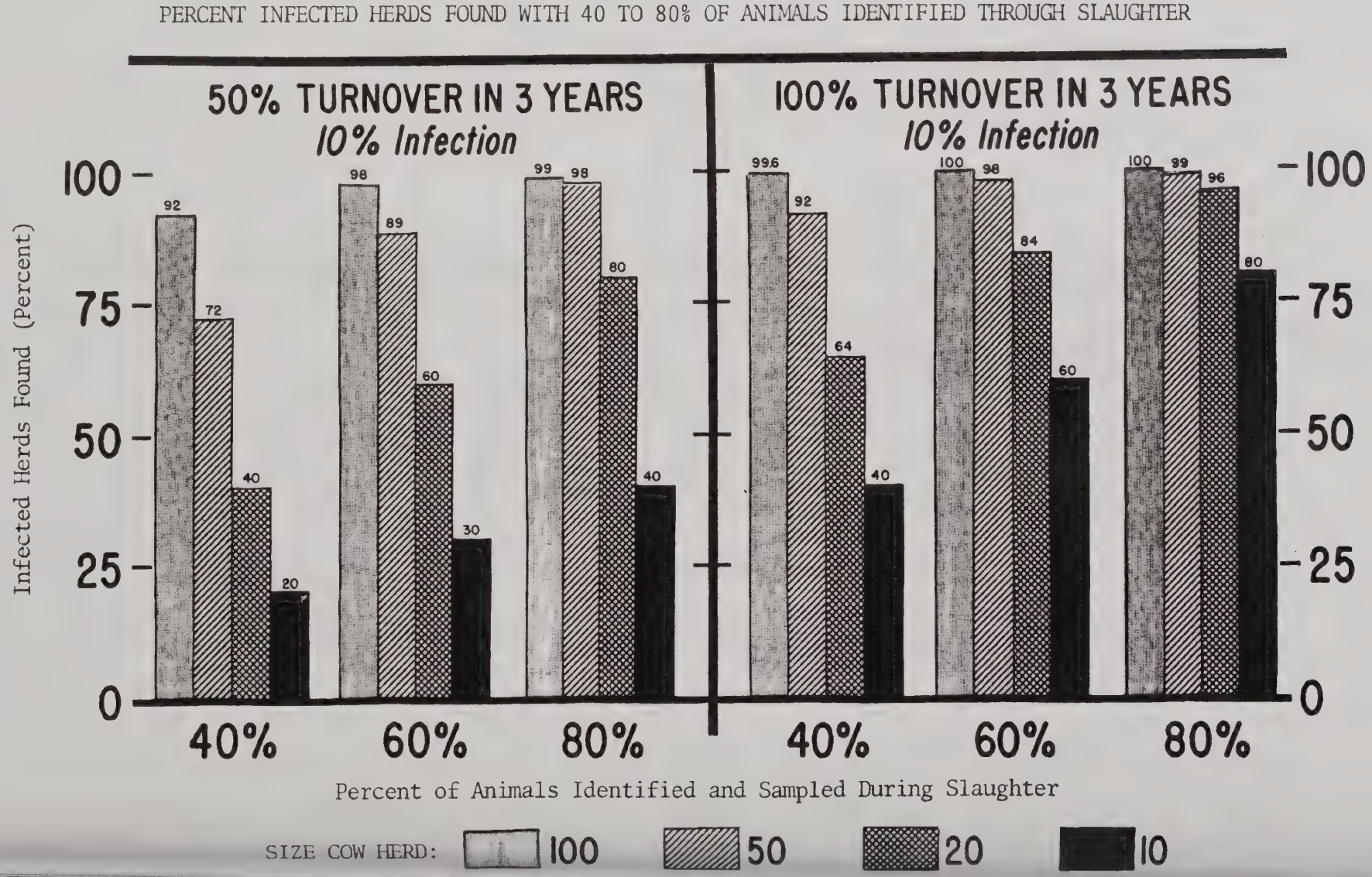


FIGURE 2 ---- PERCENT ANIMAL EXAMINATION AND SAMPLING REQUIRED

TO FIND 75 TO 95% OF INFECTED HERDS IN A 3-YEAR PERIOD

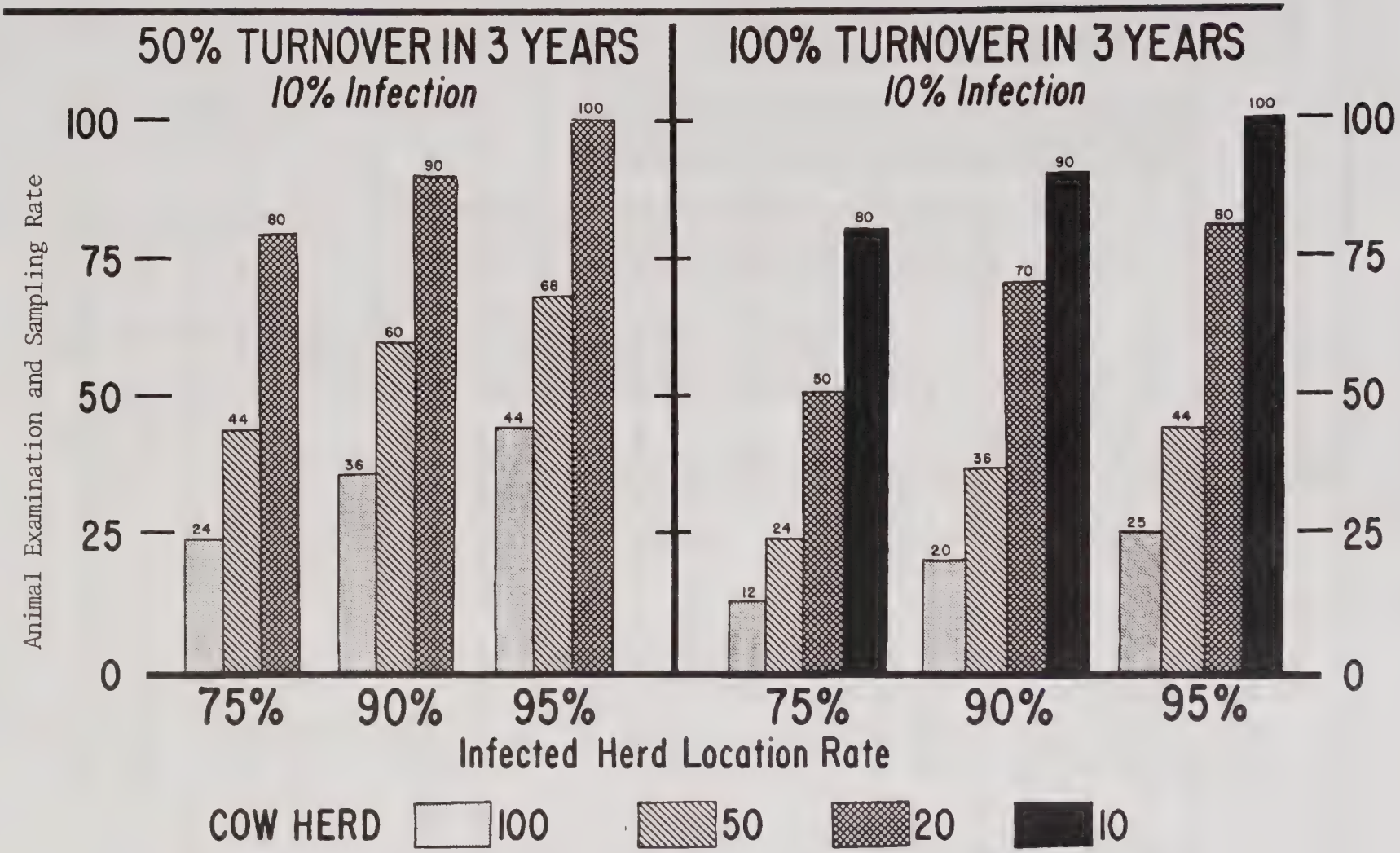
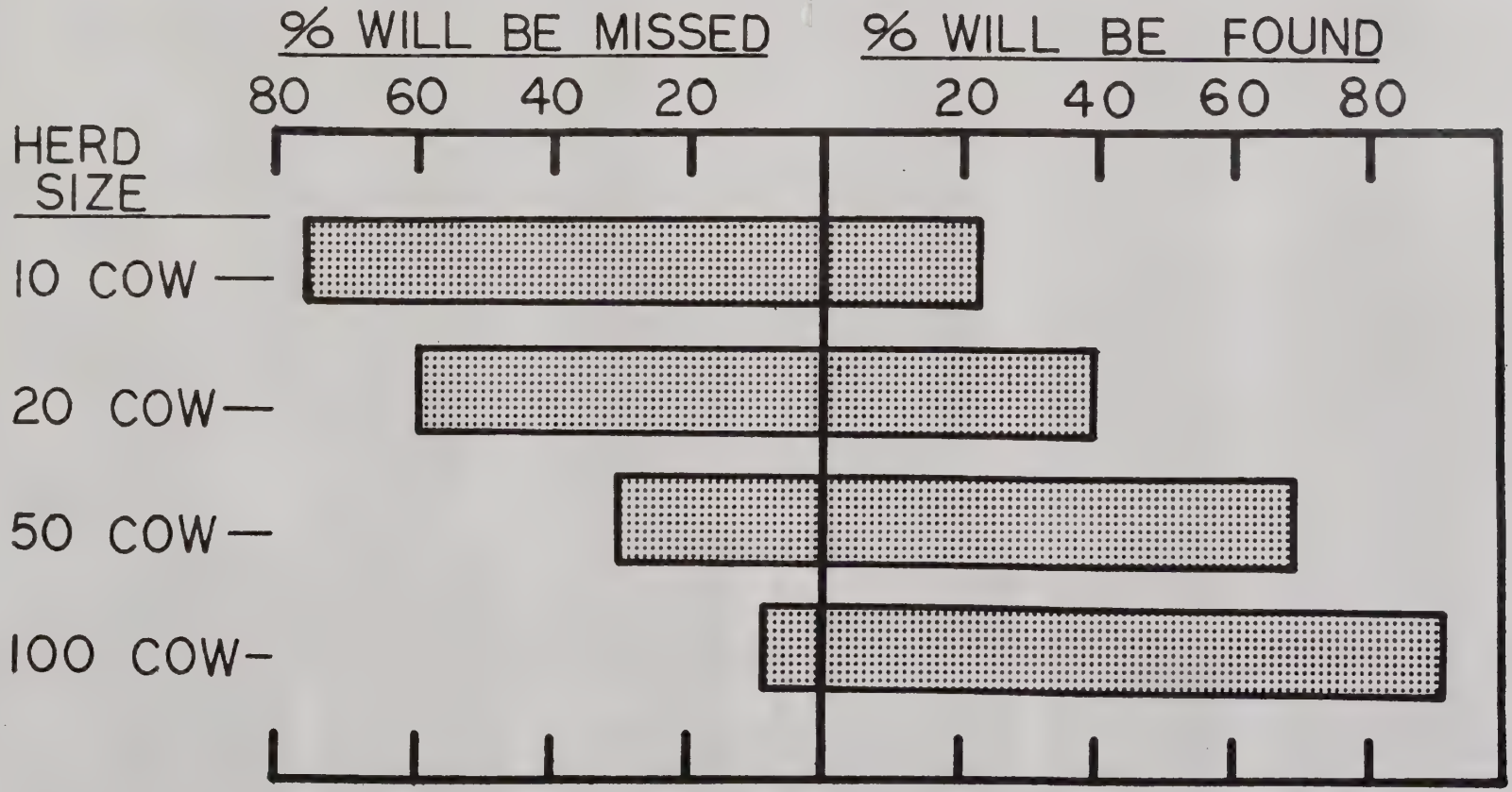


FIG.-3 PROBABILITY OF FINDING INFECTED HERDS UNDER MCT (40% COVERAGE)



BASED ON 10% HERD INFECTION BEEF TYPE CATTLE 100% TURN-OVER IN 6 YEARS.

REGULATORY STATISTICS PART V

CONSIDERATIONS ABOUT THE PROPER DESIGN AND ANALYSIS OF FIELD STUDIES AND THE PROPER RECORDING OF PERTINENT DATA

1. Purposes behind field studies. We might want to conduct a field study to find out the affect of vaccination with strain 19 vaccine and injection with 45/20 bacterin upon reducing or preventing Brucellosis or we might want to find out what relationship type of raw material and storage time has on Salmonella contamination of animal protein. In either case we are concerned with comparing rates. In the one case we compare disease rates and in the other we compare contamination rates. Any regulatory veterinarian might be faced at any time with the situation where he might have to compare rates.

   If only one set of factors are being compared such as vaccinated versus not vaccinated or type of raw material the analysis is usually simple and we would use the Chi-Square test shown in Part III. However, if we are comparing two sets of factors such as those described earlier we frequently encounter problems. Two of these problems involve presence of interaction and unequal numbers of observations in the various groups. We will discuss the nature of these problems.

2. Interactions. We will discuss only the simplest type of interaction. In doing this we will use examples from two different sources. One source consisted of a dummy example which was in a laboratory magazine. The other example involves real life data. The real life example involves part of the data from the previously mentioned brucellosis study.

If we have two different factors each consisting of two main groups or treatments we have a total of 4 sub-groups. We have an interaction when the sub-group rates are not equal to the sum of the main group rate affects.

2.1 Meaning of interactions. Interaction means different things to different people. A chemist thinks of reaction between molecules while an endocrinologist thinks of the interplay between the glands. We are concerned with the meaning as involved in an experiment or epidemiological investigation. Factors are in interaction when their effects are not additive. We learned from arithmetic that 2 + 2 = 4. But in agriculture when factors are in interaction 2 + 2 may equal 12 or 20 or 0 or some other number than 4.

We can illustrate this by the following example. Suppose that in an experiment with mice that we either feed salt or do not feed salt and that we either feed phosphorous or do not feed phosphorous. We are interested in some response variable such as body weight. Figure 1 represents the experimental design that illustrates the four different treatment combinations that are possible.

Figure 1. Experimental Design.

		A	
		0	+
B	0	00	0+
	+	+0	++

Factor A is either absent, (0), or present, (+), and Factor B is either absent, (0), or present, (+). When neither factor is present we have 00 as shown in Figure 1; if A is present and B is absent we have 0+; if A is absent and B is present we have +0; and if both are present in the diet we have ++.

In our make believe experiment we are testing the effect of four different diets: (1) neither A nor B; (2) just A; (3) just B; (4) both A and B. Figure 2 shows the numerical results that would be indicative of no interaction. The verdict of "no interaction" is reached because the results are simply additive. Feeding diet A always results in an increase of 2, (2 + 2 = 4; 5 + 2 = 7), and feeding diet B always results in an increase of 3, (2 + 3 = 5; and 4 + 3 = 7).

Figure 2. No INTERACTION

		A	
	0	+	
B 0	2	4	
+	5	7	

But Figure 3 shows different results that demonstrate the consequences of interaction.

Figure 3. INTERACTION

		A	
	0	+	
B 0	2	4	
+	5	24	

A and B have their usual effect when they are supplied independently. Thus, A alone adds 2 to 2 to make 4 and B alone adds 3 to 2 to make 5. But when A and B are fed together the result is not 7 as in Figure 2, but a much larger 24. That non-additive result is the result of interaction. It will be obvious that once interaction has been found to occur that we can never describe accurately the role of diet factor A without specifying the presence or absence of B and vice versa. It is obvious that in order for workers in the field of animal diseases to understand the workings of multiple causes that the concept of interaction must be understood. The goal of investigation in animal diseases should be the detection of the multiple parameters or factors in the sphere of interest, and the measurement of the extent of the respective interactions in statistical terms.

2.2 Interaction in a vaccine trial. We shall now illustrate interaction with part of the vaccine data mentioned previously. We will not be calculating exact values of the main affects and of the interaction since the numbers of animals in the experiment were not large enough to measure these affects without large amounts of error. However, the experiment will suffice for the purpose of illustrating interaction. We shall illustrate the data both in the form that was used in the previous example and also in the common form used in reporting data. In the experiment from which the data are abstracted there were six groups of animals. These groups resulted from the presence of animals that either were or were not vaccinated with strain 19 vaccine and from being injected with either of two types of 45/20 bacterins or not being injected. The infection rates and

experimental design are shown in Figure 4. We have used only one of the 45/20 bacterin groups. Consequently we have two main vaccine groups and two main bacterin groups which results in a 2 x 2 table.

Figure 4. Bacterin experiment data. Rates of infection.

		Bacterin status	
		45/20 present	45/20 absent
Vaccine Status	Strain 19 present	0.364	0.545
	Strain 19 absent	0.400	0.833

We could examine the data by starting either with the group which had no vaccine or bacterin or with the group which has both. We shall start with the group that has neither. By starting with the group that has no artificial protection and adding the bacterin we have decreased the infection rate from 0.833 to 0.400 or by 0.433. If we went the other way and added the strain 19 vaccine we would reduce the infection rate from 0.833 to 0.545 or by a factor of 0.288. If there was no interaction and the affects were additive then by adding strain 19 to the group that had bacterin we would have 0.400 minus 0.288 or 0.112 as the infection rate with both bacterin and vaccine present. Likewise if we added bacterin to the vaccinated group we would have 0.545 minus 0.433 or 0.112. But instead of having 0.112 as the infection rate for the group with both bacterin and strain 19 we have an infection rate of 0.364.

We shall now examine the data in the form which is standard in statistical analysis. Figure 5 shows group and sub-group rates while Figure 6 shows the affects due to adding the vaccine, bacterin, and the interaction between the two.

Figure 5. Bacterin experiment rates.

	Bacterin status		
	45/20 present	45/20 absent	Average
Vaccine Status: Strain 19 present	0.364	0.545	0.4545
Vaccine Status: Strain 19 absent	0.400	0.833	0.6165
Average	0.382	0.689	0.5355

Figure 6. Bacterin experiment affects.

	Bacterin status		
	45/20 present	45/20 absent	Average
Vaccine Status: Strain 19 present	+ 0.063	- 0.063	- 0.0810
Vaccine Status: Strain 19 absent	- 0.063	+ 0.063	+ 0.0810
Average	- 0.1535	+ 0.1535	0.5355

If we take the affects shown in Figure 6 starting with the overall average rate of 0.5355, subtract the average affect 0.1535 due to presence or absence of bacterin, subtract the average affect 0.081 due to absence or presence of strain 19 and add the interaction affect 0.063 we have the infection rate for presence of both strain 19 vaccine and 45/20 bacterin. Thus 0.364 = 0.5355 + (-0.1535) + (-0.081) + (+0.063).
The average overall infection rate of 0.5355 is equal to the average of the four sub-group rates shown in Figure 5. The average affect of having strain 19 present is equal to the average of 0.364 plus 0.545 or 0.4545. The other average rates shown in Figure 5 are calculated in the same manner. The average affect for presence of strain 19 of -0.081 shown in Figure 6 is equal to 0.4545 minus 0.5355. All other average affects are

equal to the respective average rates from Figure 5 minus the overall rate of 0.5355.

It must be remembered that the affects shown in Figure 6 are only estimates based upon small numbers and that tests of significance would reveal whether the affects due to bacterin, vaccine, and the interaction between the two were significant. The main thing to be learned is that we have interaction when the sub-group rates or means cannot be calculated when you know only the main group means. Remember that all animals getting 45/20 bacterin without regard to strain 19 vaccine constitute a main group while all animals getting both 45/20 bacterin and strain 19 vaccine constitute one of the four sub-groups in this example.

3. Consequences of unequal numbers. We shall be discussing several examples in this chapter which deal with the consequences of unequal numbers in sub-groups when they are ignored in the analysis of the data. However, in this section we shall deal with an example which would have led to a complete reversal of results if the unequal numbers had been ignored. The data in this example constitute part of the data from my Ph. D. thesis. The particular experiment involved the crossing of Red Dane, Red Poll, and Milking Shorthorn cattle. We will be looking at the results of the Red Poll by Milking Shorthorn cross. There were 79 animals involved in this particular cross of Red Poll males and females with Milking Shorthorn males and females. Figure 7 shows the experimental design and the numbers of animals in each group and sub-group. Figure 8 shows the average milk production for 300 days for the various sub-groups calculated in the same manner as for Figure 5 by averaging sub-group

averages, and Figure 9 shows what the main group averages would have been by taking total milk production divided by total animals for each group.

Figure 7. Numbers of animals in each group and experimental design of crossbreeding experiment.

		Breed of Sire	
	Red Poll	Milking Shorthorn	Total
Breed of Dam			
Red Poll	24	19	43
Milking Shorthorn	10	26	36
Total	34	45	79

Figure 8. Average Milk Production for each Sub-group and Main Group on basis of averaging averages.

		Breed of Sire	
	Red Poll	Milking Shorthorn	Average
Breed of Dam			
Red Poll	3,301	6,574	4,938
Milking Shorthorn	3,663	5,487	4,575
Average	3,482	6,030	4,756

Figure 9. Average milk production for each main group computed on basis of total production divided by total animals.

		Breed of Sire	
	Red Poll	Milking Shorthorn	Average
Breed of Dam			
Red Poll			4,757
Milking Shorthorn			4,980
Average	3,407	5,946	4,853

For example the average for Red Poll dams in Figure 9 is equal to the number of animals (24) multiplied by the milk production (3,301) for Red Poll sires plus the number of animals (19) multiplied by the milk production (6,574) for Red Poll Dams by Milking Shorthorn sires divided by total number of animals and equals 4,747: or

$$(24 \times 3,301 + 19 \times 6,574) / 43 = 4,747$$

The other averages in Figure 9 are calculated in the same manner. The important thing to observe from Figures 8 and 9 is that in Figure 8 Red Poll dams have a higher average than do Milking Shorthorn Dams while direction of the difference is reversed in Figure 9 with Milking Shorthorn Dams having the higher average. The reason why this example was chosen was that this was one where the direction of the difference would have been reversed had an incorrect analysis been performed on the data.

4. Some examples of mistakes in design and analysis. There are three different studies which illustrate the importance of proper design and analysis of studies along with the proper recording of data. Two of these studies occured in the field of regulatory veterinary medicine while the third involves the evaluation of the oral diabetes drug Tolbutamide.

5. Consequences of improper coding. One of the studies involved the Animal Health Division. This particular study was well designed and thought out. It involved broiler chickens. The study consisted of three breeds of chickens, three sexes (males, females, and straight run), and four ages of slaughter. The study came to my attention about a year ago. Even though the study involved these factors of breed, sex, and age of slaughter, these three effects upon leukosis condemnations and mortality were not evaluated. It was our hope to analyze for these effects and we examined the original data. We were able to recover records of the results of slaughter. It was found that pens that were supposed to have straight run broilers might have all males or all females, it was found that pens that were supposed to have all males might have all females or vice versa. The fact that the sex was incorrectly recorded cast doubt about the proper identification of the breed. The main object lesson from this particular study is the importance to have proper identification of the factors and animals involved in the study.

6. An example of inadequate examination of data. The second of these studies was a rendering plant study that was conducted in one of the Northern states. The main object lesson to be learned from this particular study is to look at all factors that may be of importance. The second lesson is to obtain enough samples so that tests of significance will have some meaning. A third lesson and one which many times cannot be corrected in studies involving animals or epidemiology is that of the importance if possible to have equal numbers of samples in each group.

Some of the factors of interest that were recorded are as follows:

1. Type of raw material
   a. Packer plant waste
   b. Dead animals

2. Type of finished product
   a. Pressed cake
   b. Cracklings
   c. Ground Scrap
   d. Blood Meal
   e. Other

3. Storage time
   a. long
   b. short

4. Sanitation
   a. Very Good
   b. Good
   c. Fair
   d. Poor

6.1 Type of finished product and small numbers. The first thing that we will look at is type of finished product. The important thing to be learned here is to take enough samples.

Table 1. Raw material and finished product results.

	Type of raw material			
	DEAD		PACKER	
	+	-	+	-
Pressed cake	3	17	0	4*
Cracklings	2	38	4	20
Ground Scrap	2	33	6	21
Blood Meal	0	13*	3	13*
Other	0	11*	1	2*

*Indicates numbers that are too small.

The groups where an insufficient number of samples were obtained are indicated by an asterick. Since the investigator wished to learn about the effect of type of finished product upon contamination rate more samples should have been obtained.

6.2 The consequences of examining one factor at a time. A common mistake in the examination of data like this is to look at one factor at a time. The author looked at type of finished product, type of raw

material, and storage time one at a time. There are not enough observations to allow looking at all three of these factors but we shall look at type of raw material and storage time in order to show some of the consequences of inadequate examination of the data.

Table 2. Data by individual groups.

TYPE OF MATERIAL	NUMBER OF SAMPLES Positive	Negative	Total	PER CENT POSITIVE
Long-Packer	8	23	31	25.81
Long-Dead	2	20	22	9.09
Short-Packer	6	37	43	13.95
Short-Dead	5	92	97	5.15
Total	21	172	193	

Table 3. Adjusted and unadjusted percentage positive for main groups.

	Adjusted %	Unadjusted percent	Difference
Long storage	17.45	18.87	- 1.42
Short storage	9.55	7.86	+ 1.69
Packer waste	19.88	18.92	+ 0.96
Dead animals	7.12	5.88	+ 1.24

Table 4. Differences among groups, adjusted and unadjusted.

	Adjusted	Unadjusted	Increase or decrease
Long minus short storage	7.89	11.01	- 3.12
Packer minus dead animal	12.76	13.04	- 0.28

6.2.a. Adjusted and unadjusted sub-group rates. An examination of Table 3 shows that the adjusted per cent for long storage is lower and the adjusted per cent for short storage is higher than are the unadjusted per cents. The per cent is adjusted by averaging the per cents from Table 1. For example the adjusted per cent positive for long storage is the average of 25.81 and 9.09% and is equal to 17.45% or (25.81 + 9.09) / 2 = 17.45. The unadjusted per cent is merely the total positives divided by the total samples and is 18.87% or (8 + 2) / (31 + 22).

An examination of Table 4 shows that the difference between the per cent positive for long storage and short storage is reduced from 11.01% down to 7.89% by using adjusted rates. This is a change of 3.12% by adjusting. This shows the affect of unequal numbers upon differences in various conditions.

6.2.b. Distortion of the test of significance. We will now show the results of tests of significance for this data. The fact that there are unequal numbers in the various groups pose problems in the analysis. There is in my opinion no real good way of analyzing data of this sort where there are several various types of groups and the response is positive or negative, live or die, etc. The best thing to do is to try to have equal numbers in each sub-group.

Figure 10. Test of significance for raw material ignoring storage time.

	+	-	Total	
Dead	7	112	119	
Packer	14	60	74	$X^2 = 7.997$
Total	21	172	193	

Figure 11. Test of significance for storage time ignoring raw material.

	+	-	Total	
Long	10	43	53	
Short	11	129	140	$X^2 = 4.807$
Total	21	172	193	

6.2.c. Significance levels. The Chi-Square value for a two-x-two table or 1 degree of freedom that is required to be significant at the 5% level is 3.84. The 1% Chi-Square value is 6.63 while the ½% value is 7.88. If we went by the results of the above tests we would conclude that the difference in contamination rate due to type of raw material was significant at the ½% level while the difference due to storage time was significant at the 5% level. We shall show that this is an incorrect conclusion.

Figure 12. Test of significance for storage time within the packer group.

	+	-	Total	
Long	8	23	31	
Short	6	37	43	$X^2 = 1.65$
Total	14	60	74	

Figure 13. Test of significance for storage time within the dead group.

	+	-	Total	
Long	2	20	22	
Short	5	92	97	$X^2 = 0.50$
Total	7	112	119	

Figure 14. Test of significance for raw material within the long group.

	+	-	Total	
Packer	8	23	31	
Dead	2	20	22	$X^2 = 2.35$
Total	10	43	53	

Figure 15. Test of significance for raw material within the short group.

	+	-	Total	
Packer	6	37	43	
Dead	5	92	97	$X^2 = 3.19$
Total	11	129	140	

6.3 Ways of Examing data.

6.3.a. Tests of significance. The result of each of the individual tests above is insignificant. However, in the case where each is fairly close to significance such as with the tests for raw material in Figures 14 and 15 the combined results may be significant. There are various ways of testing for the combined significance. We shall show the result of one of the methods. We shall not go into the method used since the course is not designed to teach you how to be statisticians but merely some of the consequences of looking at data wrong. Besides this particular method does not permit examination of interaction.

The breakdown given in the last set in Table 5 is given with the idea of showing that there may be an interaction between storage time and raw material in that there might not be the same difference between long and short storage time for packer waste as there is for dead animal raw material. However, we can not compute this interaction in a simple manner when we have unequal numbers in the various groups.

We can see by the results in Table 5 that there is a significant difference in type of raw material but that contrary to the conclusion drawn from Figure 11 there is not a significant difference in storage time.

TABLE 5. Analysis of Storage time and type of raw material.

Source of variation	Degrees of Freedom	Chi-Square value Actual	5%	Comment
Raw material	1	7.997	3.84	misleading
Storage time within mat.	2	2.896	5.99	
Storage time	1	4.807	3.84	misleading
Raw material within time.	2	6.086	5.99	
Raw material	1	7.997		misleading
Storage time	1	4.807		misleading
Interaction	1	-1.911		incorrect
Total	3	10.893		

6.3.b. Numerical values of the various effects. We will now show a two-by-two table of the percent positive for the different sub-groups and the deviation from the averages. If the deviation in the body of the table were 0 or not significantly different from zero than it would mean that there was no interaction.

Figure 16. Sub-group averages. Deviations.

	Packer	Dead	Average	Packer	Dead	Average
Long	25.81	9.09	17.45	+ 1.98	-1.98	+ 3.95
Short	13.95	5.15	9.55	- 1.98	+1.98	- 3.95
Average	19.88	7.12	13.50	+ 6.38	-6.38	0

The deviation of + 3.95 shown for long storage time is calculated as follows: 17.45 - 13.50 = 3.95.

The deviation for long time and packer waste of + 1.98 is calculated as follows: 1.98 = 25.81 - 3.95 - 6.38 - 13.50.

6.4 Lessons to be learned. There are various lessons to be learned from this paper. One is the necessity to put all variables into an analysis. The author concluded that there were differences in contamination due to storage time and type of raw material. We can conclude from the analyses done here that there are no differences due to storage time. It might also be that there would have been no differences due to type of raw material if differences among plants and among type of finished product had been examined also. The type of analysis done here is also inadequate. We have illustrated the clouding effect of type of raw material upon differences in storage time with the unequal numbers

causing the appearance of a significant effect when there was none.

We must remember that it is important in a field study to record all variables that may have an affect upon the variable of interest.

7. Considerations About A Diabetes Study. There has been considerable controversy in the newspapers in Washington D. C. and among medical circles about the conclusions to be drawn from a study which was done to evaluate the efficacy of an oral drug for the treatment of Diabetes Mellitus. The drug was Tolbutamide. There were four groups of patients in the study. One group received a placebo. A second group received the oral drug or insulin substitute. Two other groups received insulin. One group received a standard dose of insulin. The second group received a variable dose of insulin. There were 13 different variables recorded for each patient at the start of the study. One variable was age. 41.5% of the patients receiving the placebo were over 55 years of age. 48.0% of the patients receiving the oral insulin were over 55 years of age. The older the patient the more susceptible to mortality. Table 6 shows the percent of patients in each group having various characteristics with the rank of favoritism for assignment of patients with less risk of mortality. The controversy was over whether or not the patients receiving oral insulin were more apt to die of cardivascular causes.

Table 6. Rank of favoritism of assignment of patients.

	Placebo	Tolbutamide	Standard Insulin	Variable Insulin
Age - 55+	41.5 (1)	48.0 (4)	46.2 (3)	46.1 (2)
Sex - Female	69.3 (3)	69.1 (4)	72.9 (2)	77.5 (1)
Race - White	50.2 (2)	52.9 (3)	49.0 (1)	59.3 (4)
Hypertension present	36.8 (4)	30.2 (2)	30.9 (3)	28.1 (1)
Digitalis use	4.5 (1)	7.6 (4)	5.8 (3)	5.0 (2)
Angina Pectoris	5.0 (2)	7.0 (3)	7.7 (4)	3.5 (1)
ECG abnormality	3.0 (1)	4.0 (2½)	5.3 (4)	4.0 (2½)
Cholesterol + 300 mg.	8.6 (1)	15.1 (3)	16.4 (4)	13.4 (2)
Glucose + 110 mg.	63.5 (1)	72.1 (4)	63.6 (2)	68.0 (3)
Relative body weight +1.25	52.7 (4)	58.8 (1)	57.1 (2)	53.9 (3)
Visual acuity +20/200	4.3 (1)	5.2 (2)	6.1 (4)	5.8 (3)
Serum creatine +1.5 mg.	2.6 (4)	2.5 (3)	1.9 (1)	2.0 (2)
arterial calcification	14.3 (1)	19.7 (4)	17.2 (3)	15.9 (2)
Total of ranks.	26	39½	36	28½

Number in parentheses is rank of favoritism.

Number not in parentheses is per cent of patients in a treatment group having particular characteristic such as age over 55.

In the comparison of the placebo and tolbutamide groups the placebo group was favored with respect to 7 variables affecting mortality. The distribution was about neutral for sex, ECG, visual acuity, and creatine level. The tolbutamide group was favored with respect to Hypertension and relative weight.

7.1 Tests of significance and adjustment of data for arterial calcification. We will look at the raw data and at the breakdown for the variable which favored the placebo most over the tolbutamide in terms of changing the significance, namely arterial calcification.

Figure 17. Analysis of raw data and adjustment for arterial calcification.

RAW DATA

	Placebo	Tolbutamide	Total
Die	10	26	36
Living	195	178	373
Total	205	204	409

$X^2 = 7.884$

PERCENT DEATHS BY PRESENCE OF ARTERIAL CALCIFICATION

	Placebo	Tolbutamide	ISTD	IVAR
Absent	9.2 (174)	10.1 (159)	5.4 (168)	7.9 (164)
Present	17.2 (29)	33.3 (39)	31.4 (35)	16.1 (31)
Total patients	(203)	(198)	(203)	(195)

ADJUSTED DATA

	Placebo	Tolbutamide	Total
Die	13.35	24.355	37.705
Living	187.15	176.145	363.295
Total	200.5	200.5	401

$X^2 = 3.545$

The adjustment that has been done with respect to arterial calcification does not give the true picture. It merely indicates that the difference due to type of treatment probably is not as great as the raw data would indicate. There was also data on total mortality. The differences among

groups for total mortality do not appear to be as great as those for cardiovascular mortality.

7.2 Randomization is not a cureall.

The investigators took the attitude that since the patients were assigned to the different groups randomly and since tests for randomness did not indicate a significant departure from randomness that the variables that affect mortality such as age, etc., did not affect differences among the different treatment groups.

7.3 Lessons to be learned.

There are several things to be learned from this study. One is that randomization is not a cure all. A second is that it is necessary to look at all factors jointly. Otherwise, the unequal numbers can cause affects to be apparent which are not real. Even if there had been equal numbers of patients in each group with respect to the 13 variables affecting mortality the data should have been examined for all these factors jointly. The reason for this is that there might be certain combinations of age, sex, race, etc., that would require as the best drug the oral drug while certain other combinations might require as the best drug the regular insulin or the variable insulin.